

**Autoévaluation des connaissances partielles chez l'élève  
et mesure des manifestations métacognitives liées à son réalisme**

Serge Boulé

Thèse soumise à la  
Faculté des études supérieures et postdoctorales  
dans le cadre des exigences  
du programme de doctorat en philosophie en éducation

Mesure et évaluation

Faculté d'éducation

Université d'Ottawa

## Résumé

Dans le cadre de l'évaluation diagnostique, l'enseignant cherche à identifier les forces et les faiblesses des élèves ainsi qu'à s'informer sur leurs capacités à apprendre et à réguler leur apprentissage par eux-mêmes. Les situations de connaissances partielles et les facteurs cognitifs et métacognitifs peuvent influencer la qualité des apprentissages. Cette étude vérifie comment l'usage du degré de certitude peut contribuer à l'amélioration de l'exercice de l'évaluation diagnostique en tenant compte de ce type de situations.

Une tâche qui fait appel à des capacités cognitives éloignées de la capacité d'assimilation de l'élève peut engendrer de l'incertitude. La question est de savoir si le réalisme de l'élève (Leclercq et Poumay, 2005), établi en fonction du degré de certitude qu'il exprime envers sa performance, en est affecté. Il est également probable que la propension qu'ont les élèves à surestimer ou à sous-estimer leurs capacités puisse varier en fonction de certaines différences individuelles comme l'âge, le sexe et les capacités cognitives atteintes. Les élèves plus développés sur le plan cognitif ne sont-ils pas plus en mesure de prendre conscience et de gérer non seulement leur niveau de connaissance, mais aussi leur méconnaissance ou leur ignorance? Le niveau cognitif atteint par l'élève peut-il être considéré comme un prédicteur fiable de réalisme? Si tel est le cas, quel niveau cognitif doit être atteint par l'élève pour lui permettre des autoévaluations réalistes et fiables? Il s'agit de thèmes liés à la métacognition.

Plus de six cents élèves des deuxième, cinquième et huitième années ont été soumis à une épreuve du domaine spatio-temporel (Figures graduées de Noelting, 1980) et à une épreuve du domaine logico-mathématique (Concentrations de liquides, Noelting, 1980). Pour chaque tâche,

l'élève devait exprimer son degré de certitude par rapport à sa performance. Les tâches des épreuves variaient en matière d'exigence cognitive et de domaine cognitif, spatio-temporel ou logico-mathématique.

En ce qui concerne les stades cognitifs minimums requis pour que les degrés de certitude soient considérés utiles, le stade requis chez les garçons est un peu plus élevé que chez les filles dans les deux domaines. Il apparaît que les élèves en deuxième année commencent à être suffisamment réalistes pour les tâches du domaine spatio-temporel. Du côté du domaine logico-mathématique, il semble que le réalisme prenne plus de temps à se développer.

Les résultats soulignent une différence entre le réalisme moyen des filles et des garçons dans les tâches du domaine spatio-temporel. Les résultats indiquent aussi que le réalisme de l'élève varie en fonction du niveau scolaire. En fait, les résultats montrent une tendance baissière ou un plafonnement du réalisme de l'élève avec l'âge. Pour les tâches du domaine logico-mathématique, nous constatons aussi des différences de réalisme entre les garçons et les filles ainsi qu'entre les groupes d'âge. Il y a cependant une tendance haussière et quasi linéaire avec l'âge. Notons que la différence de réalisme entre les garçons et filles s'observe surtout chez les élèves plus jeunes.

En gros, le réalisme des élèves est plus faible pour les tâches du domaine logico-mathématique. Le phénomène se présente tant chez les filles que chez les garçons. Le réalisme des élèves serait donc dépendant du domaine. Notons que cette différence pourrait être attribuable aux modalités de *testing* des deux épreuves. L'épreuve des Figures graduées permet à l'élève de constater en temps réel et *in visu* la qualité de reproduction de ses dessins, ce qui n'est

pas possible dans le cas de l'épreuve des Concentrations portant sur le choix de la proportion ou du ratio le plus élevé.

Nous avons aussi comparé le réalisme des élèves dans les tâches qui correspondent à leur niveau opératoire ou leur sont inférieures avec le réalisme manifesté dans les tâches qui sont de niveau opératoire supérieur. L'élève est plus réaliste avec les tâches qui sont égales ou inférieures à son niveau opératoire. Le réalisme de l'élève diminue à mesure que l'exigence cognitive de la tâche s'éloigne de son niveau opératoire. Le réalisme augmente lorsque l'exigence cognitive de la tâche excède largement le niveau cognitif atteint par l'élève.

Nous concluons que l'étude du réalisme renseigne sur le développement des capacités métacognitives de l'élève et permet de mieux comprendre le rôle que celles-ci peuvent jouer dans la régulation de l'enseignement des élèves de différents niveaux scolaires.

Mots-clés :

Autoévaluation, autorégulation, degré de certitude, connaissance partielle, métacognition, évaluation diagnostique

## Table des matières

Liste des figures .....	8
Liste des tableaux .....	12
INTRODUCTION .....	17
PROBLÉMATIQUE .....	19
RECENSION DES ÉCRITS .....	22
L'évaluation diagnostique par rapport à l'évaluation formative .....	22
Historique des questionnaires à choix multiples et des degrés de certitude .....	23
Variation du réalisme et connaissance partielle .....	24
La métacognition.....	27
Les degrés de certitude .....	29
Le réalisme .....	33
Normes de réalisme.....	34
Difficulté des items .....	35
Développement cognitif .....	37
L'âge de l'élève.....	40
CADRE CONCEPTUEL .....	43
QUESTIONS DE RECHERCHE .....	53

LA MÉTHODE .....	55
Les participants .....	56
Les instruments .....	57
Épreuve des Figures graduées .....	59
Épreuve des Concentrations .....	60
Domaines et stades développementaux .....	62
Propriétés métriques des instruments .....	62
La procédure .....	63
Le plan d'analyse des résultats .....	66
Le calcul du réalisme et de l'indice de réalisation des prédictions .....	69
RÉSULTATS .....	71
La fidélité interjuge .....	72
Les cas aberrants .....	73
Les coefficients de reproductibilité .....	77
Les distributions de l'ensemble des variables indépendantes et dépendantes .....	79
Indice de réalisation des prédictions .....	79
Stades cognitifs atteints .....	81
Réalisme des élèves .....	82

VÉRIFICATION DES QUESTIONS DE RECHERCHE .....	88
Postulats .....	88
Liens entre les variables – régression multiple .....	91
Liens entre les variables – corrélation entre les variables dépendantes .....	94
Effet du stade cognitif atteint sur le réalisme des élèves .....	98
Réalisme des élèves en fonction du niveau scolaire et du sexe.....	102
Épreuve des Figures graduées .....	108
Épreuve des Concentrations .....	111
Réalisme des élèves et domaine cognitif : comparaison entre les garçons et les filles .....	114
Réalisme selon l'écart entre exigence cognitive de la tâche et stade atteint .....	117
Réalisme de l'élève dans les cas extrêmes d'écart entre l'exigence cognitive de la tâche et le niveau de développement cognitif atteint .....	126
DISCUSSION .....	133
LIMITES ET RECOMMANDATIONS .....	143
RÉFÉRENCES .....	151

## Liste des figures

Figure 1 – Distribution spectrale des réponses .....	26
Figure 2 – Exemple de modalité d’expression du degré de certitude .....	32
Figure 3 – Algorithme de modélisation du processus d’interprétation - cas de surestimation .....	48
Figure 4 -- Réalisme de l’élève en fonction de l’écart entre le niveau d’exigence de la tâche et le niveau cognitif atteint par l’élève .....	50
Figure 5 – Distribution du réalisme des élèves pour l’épreuve des Figures graduées .....	75
Figure 6 – Distribution du réalisme des élèves pour l’épreuve des Concentrations .....	76
Figure 7 – Indice de réalisation des prédictions de l’item pour l’épreuve des Figures graduées .....	80
Figure 8 – Indice de réalisation des prédictions de l’item pour l’épreuve des Concentrations .....	81
Figure 9 – Distribution du stade atteint par les élèves selon l’épreuve des Figures graduées .....	83
Figure 10 – Distribution du stade atteint selon l’épreuve des Concentrations .....	84
Figure 11 – Distribution du réalisme pour l’ensemble des élèves avec l’épreuve des Figures graduées .....	86

Figure 12 – Distribution du réalisme pour l'ensemble des élèves avec l'épreuve des Concentrations .....	87
Figure 13 – Relation entre le réalisme des filles et le stade cognitif atteint selon l'épreuve des Figures graduées .....	89
Figure 14 – Relation entre le réalisme des garçons et le stade cognitif atteint selon l'épreuve des Figures graduées .....	89
Figure 15 – Relation entre le réalisme des filles et le stade cognitif atteint selon l'épreuve des Concentrations .....	90
Figure 16 – Relation entre le réalisme des garçons et le stade cognitif atteint selon l'épreuve des Concentrations .....	90
Figure 17 – Relation curvilinéaire entre le réalisme des élèves en fonction des deux épreuves .....	97
Figure 18 – Relation linéaire entre le réalisme des élèves en fonction des deux épreuves..	97
Figure 19 – Réalisme moyen des filles en fonction du stade atteint selon l'épreuve des Figures graduées .....	99
Figure 20 – Réalisme moyen des garçons en fonction du stade atteint selon l'épreuve des Figures graduées .....	99

Figure 21 – Réalisme moyen des filles en fonction du stade atteint selon l'épreuve des Concentrations .....	100
Figure 22 – Réalisme moyen des garçons en fonction du stade atteint selon l'épreuve des Concentrations .....	100
Figure 23 – Réalisme moyen des élèves selon le niveau scolaire et le sexe pour l'épreuve des Figures graduées .....	111
Figure 24 – Réalisme moyen des élèves selon le niveau scolaire et le sexe pour l'épreuve des Concentrations .....	113
Figure 25 – Réalisme moyen des filles dans les deux épreuves selon le niveau scolaire ..	116
Figure 26 – Réalisme moyen des garçons dans les épreuves selon le niveau scolaire ....	117
Figure 27 – Écart du réalisme moyen des élèves de tous les niveaux scolaires et l'exigence cognitive de la tâche pour l'épreuve des Figures graduées .....	124
Figure 28 – Écart du réalisme moyen des élèves selon le niveau scolaire et l'exigence cognitive de la tâche pour l'épreuve des Concentrations .....	126
Figure 29 – Réalisme moyen des élèves du niveau IA à mesure que l'exigence cognitive de la tâche s'éloigne du niveau cognitif atteint par l'élève – épreuve des Figures graduées .....	129

Figure 30 – Réalisme moyen des élèves du niveau IB à mesure que l'exigence cognitive de la tâche s'éloigne du niveau cognitif atteint par l'élève – épreuve des Figures graduées .....	129
Figure 31 – Réalisme moyen des élèves du niveau IIA à mesure que l'exigence cognitive de la tâche s'éloigne du niveau cognitif atteint par l'élève – épreuve des Figures graduées .....	130
Figure 32 – Réalisme moyen des élèves du niveau IA à mesure que l'exigence cognitive de la tâche s'éloigne du niveau cognitif atteint par l'élève – épreuve des Concentrations .....	130
Figure 33 – Réalisme moyen des élèves du niveau IB à mesure que l'exigence cognitive de la tâche s'éloigne du niveau cognitif atteint par l'élève – épreuve des Concentrations .....	131
Figure 34 – Réalisme moyen des élèves du niveau IC à mesure que l'exigence cognitive de la tâche s'éloigne du niveau cognitif atteint par l'élève – épreuve des Concentrations .....	131
Figure 35 – Réalisme moyen des élèves du niveau IIA à mesure que l'exigence cognitive de la tâche s'éloigne du niveau cognitif atteint par l'élève – épreuve des Concentrations .....	132

### Liste des tableaux

Tableau 1 – Indices métacognitifs .....	26
Tableau 2 – Taxonomie révisée .....	30
Tableau 3 – Normes de réalisme en situation d'examen et d'auto-estimation .....	35
Tableau 4 – Tableau résumant les stades de développement cognitif piagétien .....	44
Tableau 5 - Lien entre la réponse à l'item et le degré de certitude exprimé .....	46
Tableau 6 – Effectifs de l'échantillon .....	58
Tableau 7 – Découpage des stades par catégorisation et validation effectuée par analyse factorielle exploratoire .....	61
Tableau 8 – Sommaire des variables .....	67
Tableau 9 – Coefficients de reproductibilité pour les 16 items de l'épreuve des Figures graduées .....	78
Tableau 10 – Coefficients de reproductibilité pour les 15 items de l'épreuve des Concentrations .....	79
Tableau 11 – Statistiques sur les distributions des stades cognitifs atteints selon l'épreuve des Figures graduées .....	83
Tableau 12 – Statistiques sur les distributions des stades cognitifs atteints selon l'épreuve des Concentrations .....	84

Tableau 13 – Statistiques sur les distributions du réalisme des élèves avec l'épreuve des Figures graduées .....	86
Tableau 14 – Statistiques sur les distributions du réalisme des élèves avec l'épreuve des Concentrations .....	87
Tableau 15 – Résumé du modèle proposé par l'analyse de la régression pour le réalisme ( $R_s$ ) de l'épreuve des Figures graduées .....	93
Tableau 16 – Résumé du modèle proposé par l'analyse de la régression pour le réalisme ( $R_s$ ) de l'épreuve des Concentrations .....	94
Tableau 17 – Corrélation entre les stades déterminés par les deux épreuves et le réalisme des élèves .....	96
Tableau 18 – Stades cognitifs requis pour que le réalisme représente une valeur ajoutée..	102
Tableau 19 – Statistiques descriptives sur l'analyse multivariée de la variance de réalisme pour l'épreuve des Figures graduées .....	103
Tableau 20 – Statistiques descriptives sur l'analyse multivariée de la variance de réalisme pour l'épreuve des Concentrations .....	104
Tableau 21 - Tableau résumé de l'analyse multivariée de la variance de réalisme pour les deux épreuves .....	106

Tableau 22 – Tableau résumé de l’analyse univariée de la variance pour les deux épreuves .....	107
Tableau 23 – Résultats des contrastes pour les deux épreuves .....	109
Tableau 24 – Tableau résumé de l’analyse de la variance selon un plan à deux critères de classification .....	115
Tableau 25 – Résultats des contrastes pour les deux épreuves en fonction du niveau scolaire .....	115
Tableau 26 – Statistiques descriptives sur le réalisme selon l’écart entre l’exigence cognitive de la tâche et le stade atteint - Épreuve des Figures graduées ..	119
Tableau 27 – Statistiques descriptives sur le réalisme selon l’écart entre l’exigence cognitive de la tâche et le stade atteint - Épreuve des Concentrations .....	119
Tableau 28 – Tableau résumé de l’analyse multivariée de la variance à mesures répétées sur les écarts de réalisme entre l’exigence cognitive de la tâche et le stade atteint pour l’épreuve des Figures graduées et pour l’épreuve des Concentrations .....	121
Tableau 29 – Tableau résumé de l’analyse de la variance à mesures répétées sur les écarts de réalisme entre l’exigence cognitive de la tâche et le stade atteint pour l’épreuve des Figures graduées et pour l’épreuve des Concentrations traitées séparément .....	123

Tableau 30 – Réalisme moyen des élèves à mesure que l'exigence cognitive de la tâche s'éloigne  
du niveau cognitif atteint par l'élève – épreuve des Figures graduées ..... 128

Tableau 31 – Réalisme moyen des élèves à mesure que l'exigence cognitive de la tâche s'éloigne  
du niveau cognitif atteint par l'élève – épreuve des Concentrations ..... 128

## Appendices

A – Exemple de figures graduées et de directives destinées à l'élève .....	163
B – Exemple provenant de l'épreuve des Concentrations .....	165
C – Modalité permettant l'expression du degré de certitude .....	167
D – Le calcul du réalisme .....	169

## INTRODUCTION

Dans le cadre de l'évaluation diagnostique, le personnel enseignant cherche à identifier les forces et les faiblesses des élèves ainsi qu'à s'informer sur leurs capacités à apprendre et à réguler leur apprentissage par eux-mêmes. Les situations de connaissances partielles et les facteurs cognitifs et métacognitifs peuvent influencer la qualité des apprentissages. Cette étude vérifie comment l'usage du degré de certitude et du réalisme peut contribuer à l'amélioration de l'exercice de l'évaluation diagnostique en tenant compte de ce type de situations.

Une bonne compréhension des conditions d'utilisation du degré de certitude et du réalisme avec des enfants nous permettra éventuellement de dépasser le contexte de l'évaluation diagnostique et explorer les avenues possibles d'implantation comme dans le cas de l'évaluation formative. Pour ce faire, nous devons d'abord vérifier l'applicabilité et la pertinence de l'utilisation du degré de certitude et du réalisme. Est-ce que ces manifestations métacognitives se mesurent? Sont-elles stables dans le temps ou d'un domaine cognitif à l'autre ou d'une matière à l'autre? Fonctionnent-elles de la même façon en fonction de facteurs interindividuels comme le sexe, le développement cognitif et l'âge de l'élève?

Le ministère de l'éducation de l'Ontario exige que les pratiques d'évaluation en salle de classe soient rigoureuses. Un format de relevé de notes qui inclut des cotes décrivant certaines habiletés et habitudes de travail est d'ailleurs recommandé par le ministère. L'autorégulation fait partie de ces habiletés de travail des élèves qui doivent être évaluées. Cette recherche est pertinente pour le personnel enseignant, car elle tente de faire la lumière sur des stratégies permettant de mesurer des manifestations métacognitives qui sont liées au réalisme des élèves lorsqu'ils s'autoévaluent. En fait, il nous apparaît que l'élève doit être minimalement réaliste

lorsqu'il s'autoévalue avant de pouvoir réguler son propre apprentissage. Le personnel enseignant ainsi que le programme de formation des enseignantes et des enseignants peuvent favoriser la progression des élèves dans l'apprentissage de la prise de conscience de leur propre processus cognitifs.

Cette recherche explore ces concepts ainsi que les questions qui s'y rattachent. Dans certains cas, nous observons des phénomènes et répondons au moins partiellement à certaines questions. Des pistes de recherches et des devis alternatifs sont également proposés.

## PROBLÉMATIQUE

Le sujet traité dans cette recherche s'inscrit dans le cadre de l'évaluation diagnostique en éducation. La plupart des instruments diagnostiques utilisés en éducation tels que, entre autres, le test de la Trousse d'évaluation en lecture GB+, le Test de rendement pour francophones (TRF), le test *Évaluation de lecture (DRA)* ainsi que le *Test d'habiletés scolaires Otis-Lennon (OLSAT)*, sont administrés à des périodes précises de l'année scolaire pour produire un bilan de l'état des apprentissages des élèves ou de leur aptitude à apprendre. Généralement, ces instruments ne permettent pas de fournir aux élèves et aux enseignants de l'information utile concernant le niveau d'autonomie des élèves, leur capacité à autoréguler leur apprentissage ainsi que leur connaissance de leurs propres processus cognitifs qui fait partie de ce qu'on appelle la métacognition.

Or, le personnel enseignant cherche effectivement à identifier les forces et les faiblesses de ses élèves ainsi qu'à s'informer sur leurs capacités à apprendre et à réguler leur apprentissage par eux-mêmes. Il lui est possible de recourir à l'interprétation des scores obtenus qui peut tenir compte des situations de connaissances partielles des élèves puisqu'ils n'ont pas toujours ni une connaissance absolue ni une certitude absolue de leurs réponses dans un contexte de *testing*. Cette étude vérifie comment l'usage du degré de certitude peut contribuer à améliorer l'évaluation diagnostique en tenant compte de ces types de situations de connaissances partielles.

La problématique de cette recherche s'inscrit dans le domaine du réalisme et de la certitude exprimée par les élèves en situation d'évaluation. Il s'agit d'une recherche fondamentale visant à identifier les conditions nécessaires pour rendre l'utilisation de la mesure

du réalisme pertinente et propice au diagnostic des apprentissages et des stratégies d'apprentissage des élèves.

L'évaluation diagnostique vise l'identification des forces et des faiblesses des élèves à un moment précis de leur apprentissage. Sur une base collective, l'évaluation diagnostique fournit des repères pédagogiques permettant de réguler l'enseignement. Sur une base individuelle, elle permet d'identifier les acquis et les difficultés éventuelles des apprenants. Selon Cardinet (1988), l'évaluation diagnostique permet d'analyser les problèmes particuliers de chaque élève et de remédier à la situation.

Cette recherche vise à enrichir l'évaluation diagnostique en prenant en compte le réalisme des élèves et leurs capacités à s'autoévaluer. Ceci permet de tenir compte du degré d'autonomie et de prise de conscience de l'élève, et peut représenter une valeur ajoutée utile à la prise de décision par le personnel enseignant et l'élève lui-même. Voilà pourquoi l'étude du degré de certitude et du réalisme à des fins d'évaluation diagnostique présente un grand intérêt pour l'examen du développement des opérations cognitives et des stratégies métacognitives chez l'élève.

Cette recherche s'interroge donc sur des facteurs qui peuvent influencer l'évaluation du réalisme de l'élève tels que la complexité de la tâche, l'exigence cognitive de la tâche, l'écart entre l'exigence cognitive de la tâche et le niveau développemental atteint par l'élève, le domaine cognitif de la tâche, le sexe de l'élève. Nous examinons également l'âge à partir duquel l'élève peut autoestimer *a posteriori* la qualité de sa performance avec réalisme.

L'étude des degrés de certitude soulève également une autre problématique de taille. Qu'il s'agisse de différencier les élèves entre eux dans un contexte d'évaluation sommative ou de différencier les compétences d'un élève à des fins diagnostiques, il faut composer avec les problèmes causés par les réponses au hasard qui nuisent à la capacité de différencier les élèves. On parle alors de correction due au hasard. Dans une perspective d'évaluation sommative, la correction due au hasard permet de juger les scores obtenus en tentant d'établir le score vrai avec davantage de précision. Dans un contexte d'évaluation diagnostique qui tient compte du réalisme des élèves, l'utilisation des degrés de certitude peut entre autres contribuer à minimiser l'impact des réponses au hasard et tenir compte d'états de connaissances partielles (voir section suivante).

Pour cela, nous abordons cette étude avec une recension des écrits sur les différents aspects de la problématique. Après un bref historique portant sur les questionnaires à choix multiples et les degrés de certitude, l'état de la question est présenté sur les liens entre le réalisme, le sexe de l'élève, l'âge de l'élève, le niveau de développement cognitif de l'élève et l'exigence cognitive évoqué par les tâches. Ensuite, le cadre conceptuel explique comment est interprété l'ensemble de ces concepts et enchaîne avec les questions de recherche. La section de la méthode révèle la stratégie employée qui permet d'obtenir les données nécessaires traitées dans la section des résultats. Cette étude se termine par des considérations liées aux limites de l'étude et par une discussion qui résume les réponses aux questions de recherche.

## RECENSION DES ÉCRITS

### **L'évaluation diagnostique par rapport à l'évaluation formative**

Selon Legendre (2005), l'évaluation formative se définit ainsi :

Un processus d'évaluation continue ayant pour objet d'assurer la progression de chaque individu dans une démarche d'apprentissage ou le rythme de cette progression pour apporter des améliorations ou des correctifs. (p. 85)

Scallon (2000) rapporte que l'appellation « évaluation formative » trouve son origine chez Scriven (1967) qui proposait de l'appliquer dans un contexte d'évaluation de programme. Les concepteurs de programme utilisent des méthodes d'évaluation dites formatives en cours d'implantation dans une perspective d'amélioration continue jusqu'au moment de l'évaluation sommative. Scriven s'est inspiré de Cronbach (1963) qui argumentait en faveur de l'amélioration de l'enseignement, la détermination des besoins des élèves et la régulation administrative. Bloom, Hastings et Madaus (1971) transposent les idées de Scriven dans le contexte de l'apprentissage scolaire. Dès lors, la régulation des apprentissages est centrale en évaluation formative. On se donne des outils pour remédier rapidement à toutes difficultés individuelles et collectives qu'on rencontre durant le déroulement d'une séquence d'apprentissages.

Scallon (2000) et Bloom, Madaus et Hastings (1981) font état d'une certaine confusion à l'égard de la définition de l'évaluation diagnostique et de ce qui la démarque de l'évaluation formative. En fait, même l'appellation « évaluation diagnostique » n'est pas universelle. Cardinet (1977) parle de fonction prédictive de l'évaluation alors que De Ketele (1993), Allal (1991) et Hadji (1997) parlent de fonction pronostique de l'évaluation. Ce qui la différencie de l'évaluation

formative se situe surtout dans l'intention et le moment de l'évaluation. Scallon considère que ce type d'évaluation peut avoir une fonction autre que pédagogique comme dans les cas du classement des élèves, de leur aiguillage vers des programmes ou encore de l'orientation scolaire et professionnelle. Grégoire (1999) souligne néanmoins que le diagnostic doit s'appuyer sur une connaissance des processus mentaux. Notons que les résultats d'une évaluation diagnostique ne sont pas nécessairement communiqués aux élèves, ne sont pas nécessairement liés à un contenu d'apprentissage récemment acquis et ne visent pas essentiellement à informer les élèves pour favoriser la régulation de leur apprentissage; ce qui est tout à l'opposé de l'évaluation formative.

### **Historique des questionnaires à choix multiples (QCM) et des degrés de certitude**

Spearman (1907) jette les fondements de ce qui est connu comme la *Théorie classique des scores* (TCS). Dès lors, les tests sont composés de plusieurs items corrigés de façon dichotomique : les réponses sont bonnes ou mauvaises. En évaluation, une des problématiques relevées concerne la question des crédits partiels. Il y a des situations de connaissance complète comme celle où seule une réponse exacte est possible. Ce sont souvent des questions de connaissance selon la taxonomie des objectifs cognitifs de Bloom ou de la taxonomie révisée d'Anderson, Krathwohl, Airasian, Cruikshank, Mayer, Pintrich, Raths & Wittrock (2001). Par contre, il existe d'autres situations qui consistent à déduire ou à donner une approximation comme celle d'une estimation de la distance entre deux villes. Il s'agit alors de connaissance partielle où la réponse se situe sur une échelle continue. Plus l'estimation est précise, plus elle est un bon indicateur d'un niveau élevé de capacité. Les items corrigés de façon dichotomique ne permettent pas directement de rendre compte de ce type de situation.

Notons qu'en évaluation sommative, les niveaux de connaissances partielles peuvent être traduits en crédits partiels. Dans un contexte où on utilise des items et des tâches nécessitant des réponses courtes et élaborées, l'évaluation du degré de connaissances partielles peut se quantifier à l'aide d'échelles descriptives d'appréciation ou de grilles de notation. Quant aux questionnaires à choix multiples (QCM), ceux-ci présentent des choix de réponses qui sont plus pertinents que d'autres et leur pondération peut refléter cette réalité. Pour composer avec les questions de connaissances et de crédits partiels, il existe différentes méthodes; à titre d'exemple, Haladyna (2004) propose sa méthode de « *polyscoring* » qui consiste à demander à l'élève de mettre les choix de réponses en ordre de plausibilité. Plusieurs autres propositions s'inscrivent dans les modèles multidimensionnels de réponses à l'item (Kelderman, 1996; Muraki, Carlson & Bolt, 1996; Yao & Schwarz, 2005). Cela peut aussi être illustré au moyen d'un exemple de question portant sur une opération d'arithmétique; l'élève qui ne respecte pas l'ordre des opérations s'éloigne davantage de la bonne réponse que celui qui effectue une erreur de calcul; de plus, le cas est nettement aggravé si l'élève ne respecte pas l'ordre des opérations en exprimant un degré de certitude élevé par rapport à la qualité de sa réponse.

### **Variation du réalisme et connaissance partielle**

La mesure du réalisme est obtenue en vérifiant la concordance entre le degré de certitude exprimé à l'égard de la réponse choisie et la qualité de cette réponse. L'élève réaliste a tendance à anticiper une bonne réponse lorsque celle-ci est effectivement bonne. Il anticipe également une mauvaise réponse lorsqu'elle est mauvaise.

Legendre (2005) définit le réalisme comme le résultat de la prise en compte de toutes les données objectives d'une situation afin de les apprécier avec justesse. Leclercq et Poumay (2005)

définissent les degrés de certitude comme des « jugements, analyses et/ou régulations observables ou non observables effectués par l'élève sur ses propres performances (processus ou produits d'apprentissages), dans des situations de PRÉ, PER ou POST performance ». La perception correcte par un élève de la qualité de ses réponses fait partie de la notion que De Landsheere (1979) appelle une image de soi réaliste, complète et stable.

Dans un contexte de connaissance partielle, Hunt (1993) classe les états de la connaissance selon les catégories suivantes : connaissance, ignorance et méconnaissance. Les travaux les plus récents de Leclercq, Rinaldi & Emould (2003), Gilles (2005), Leclercq (2011), Peraya, Grenon et Jenni (2011) visent l'opérationnalisation de ces concepts telle qu'illustrée dans la Figure 1. Il s'agit des « états de connaissance » correspondant à différents niveaux de réalisme. Leclercq et al. les situent sur une échelle allant de « connaissances nuisibles » (réponse incorrecte avec certitude élevée), qui sont des cas de surestimation de la qualité des réponses ou de fausses représentations, à « connaissances utilisables » (réponse correcte avec certitude élevée) qui sont des cas de réalisme élevé. En fait, Leclercq et Poumay (2003) utilisent plusieurs indices métacognitifs pour qualifier les différents degrés de réalisme tels que présentés dans le Tableau 1. Notons que l'échelle de réalisme utilisée consiste en un coefficient de corrélation bisériale qui est calculé sur la base des étudiants les plus réalistes. Gilles (2002) parle de « qualité spectrale des réponses ». Leclercq (2003) explique que, dans un contexte d'apprentissage, on ne passe pas d'une connaissance nulle à une connaissance totale, mais à différents degrés de connaissance partielle. L'évaluation diagnostique des connaissances partielles fournit des renseignements utiles sur la capacité d'apprendre de l'élève par lui-même. Elle fait plus que corriger les effets du hasard. Une question demeure dans un contexte d'évaluation diagnostique.

Considérant que les stades piagétiens s'acquièrent selon un ordre invariant, peut-on passer d'un état de connaissance partielle telle que « l'ignorance avouée » à « connaissance utilisable » sans passer par « connaissance incertaine »?

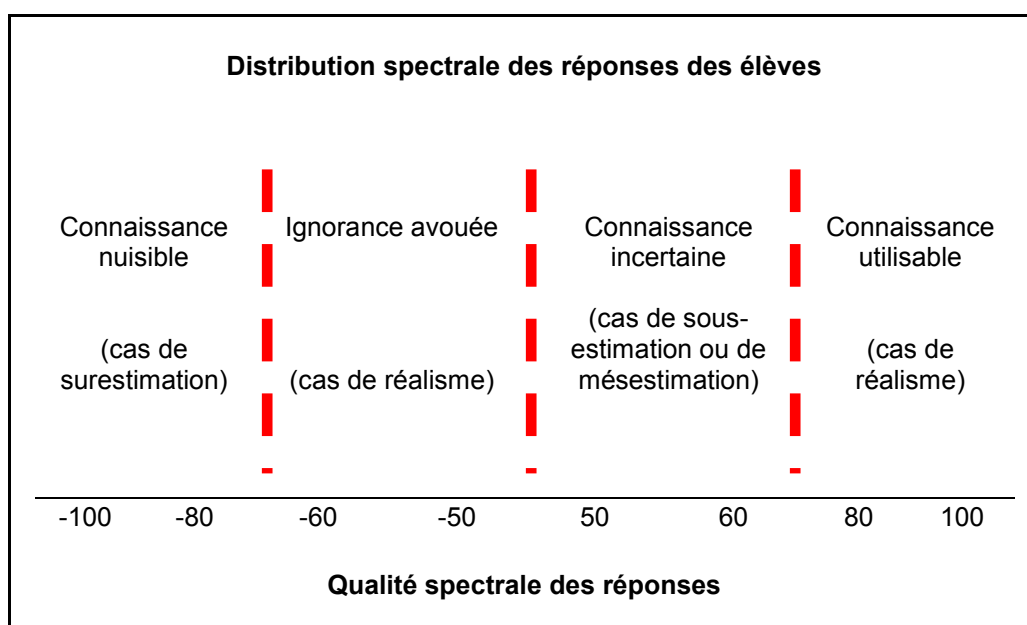


Figure 1. Distribution spectrale des réponses (adapté de Leclercq, Rinaldi & Ernould, 2003).

Tableau 1

Indices métacognitifs (adapté de Leclercq et Poumay, 2003, p.185)

	Réponses incorrectes		Réponses correctes	
	> 50 %	< 50 %	< 50 %	> 50 %
Certitude	> 50 %	< 50 %	< 50 %	> 50 %
Statut cognitif	Méprise	Mé-connaissance	Mi-connaissance	Maîtrise
Statut métacognitif	Imprudence	Prudence	Manque de confiance	Confiance
Qualificatif	Nuisible	Inutilisable	Améliorable	Assurée

Il est probable que les recherches de Leclercq et de ses collaborateurs ainsi que la plupart des outils utilisés auprès d'étudiants adultes de niveau universitaire s'appliquent aux élèves également. Notons cependant que ces étudiants adultes ont vraisemblablement déjà atteint des niveaux avancés sur le plan cognitif. De par leur échantillon de circonstance, leurs recherches ne prennent pas en considération toutes variables liées au développement cognitif et au développement cognitif minimum requis pour être en mesure de s'autoévaluer. Pour être utilisable auprès d'élèves plus jeunes, il est pertinent d'étudier l'évolution des connaissances partielles telle qu'elle s'inscrit à l'intérieur du développement cognitif. La distribution spectrale des réponses des élèves (Leclercq, Rinaldi & Ernould; 2003) et les indices métacognitifs associés (Leclercq & Poumay; 2003) ne sont pas nécessairement pertinents et représentatifs à tous les niveaux de développement cognitif. Par exemple, la méconnaissance chez un enfant ou un élève n'a pas nécessairement la même signification que chez un adulte. Dans le cas d'un élève qu'on présume attentif, l'indice de réalisme peut être biaisé par un manque de développement de l'élève sur le plan cognitif. Pour les adultes, la méconnaissance est peut-être simplement de la méconnaissance.

Il est plus probable que l'élève sera réaliste lorsque les conditions gagnantes sont réunies. Par exemple, si l'exigence de la tâche requise est en lien avec sa capacité cognitive, l'élève a plus de chance d'être réaliste par rapport à la qualité de ses performances.

### **La métacognition**

Lorsque l'élève formule une réponse, l'élève attentif porte un jugement sur l'exactitude de ses choix et prend une décision. Ce jugement est généralement fondé sur une connaissance de soi et de l'objet d'apprentissage et résulte d'un processus d'ordre métacognitif. Puisque ce

jugement découle d'un processus métacognitif, il est nécessaire de bien comprendre le processus sous-jacent à cette prise de décision et à la connaissance de l'élève de ce processus. Les liens entre l'autoévaluation, l'utilisation des degrés de certitude, le réalisme et la métacognition sont tissés par Leclercq (1993) ainsi que Leclercq et Poumay (2009).

Jans et Leclercq (1997) expriment la nécessité d'établir ou de s'aligner avec une forme de taxonomie du « réalisme métacognitif ». Celle-ci clarifierait la question et faciliterait les recherches appliquées. Dans un contexte de *testing* diagnostique, une telle taxonomie faciliterait également l'élaboration de tableaux de spécification en permettant de décomposer les tâches ou les items en fonction de critères d'ordre métacognitifs.

La taxonomie d'Anderson, Krathwohl, Airasian, Cruikshank, Mayer, Pintrich, Raths, et Wittrock (2001), contrairement à la taxonomie originale de Bloom, fournit un cadre potentiellement utile pour rendre compte de l'expression du degré de certitude en situation d'évaluation et d'autoévaluation puisqu'elle incorpore les concepts de la métacognition. Pintrich (2002) décrit la connaissance de soi (« Self-knowledge ») comme l'un des trois types de connaissance métacognitive (« Metacognitive knowledge »). L'habileté à évaluer sa propre connaissance est une composante importante de la connaissance de soi. Nous pouvons étendre cette affirmation en postulant que le réalisme de l'élève est une forme de connaissance métacognitive. Le réalisme de l'étudiant cadrerait ainsi dans la taxonomie révisée d'Anderson et al. (2001) sous la rubrique « Connaissance métacognitive » des dimensions de la connaissance (« Knowledge dimension »). Notons que cette taxonomie révisée inclut plus que des niveaux hiérarchisés de processus cognitifs. Nous trouvons des dimensions pour chaque niveau; chaque dimension étant ventilée en rubriques. Selon Anderson et al., le contrôle métacognitif et le

processus d'autorégulation sont les processus cognitifs que les élèves utilisent pour surveiller, contrôler et réguler leur cognition et leur apprentissage. Ils s'intègrent dans les six catégories de processus cognitif dans la taxonomie révisée. Nous postulons qu'il en est de même pour le réalisme. L'évaluation diagnostique du réalisme doit prendre en compte les six catégories de processus cognitif. Le niveau taxonomique des tâches d'évaluation peut affecter le réalisme des élèves. Il est probablement plus facile de s'autoévaluer avec réalisme sur une tâche qui fait appel à la mémoire que sur une tâche qui exige une analyse. Dans le deuxième cas, il est plus difficile d'évaluer la qualité de sa réponse puisqu'il est possible qu'il y ait plus d'une bonne réponse ou que chaque réponse puisse varier en matière de qualité. Le Tableau 2 présente les aspects multidimensionnels de cette nouvelle taxonomie. Un item ou une tâche pourrait se retrouver dans une ou plusieurs cellules de cette grille, un peu comme dans un tableau de spécifications.

### **Les degrés de certitude**

Le degré de certitude d'un élève est un élément qui peut être ajouté au processus d'évaluation parce qu'il permet de rendre compte des degrés variables de connaissance. Par exemple, l'élève indique, en plus de sa réponse à la question du test, son degré de certitude en utilisant une échelle de probabilités d'exactitude qui peut varier par échelons entre « *pas sûr du tout* » et « *absolument sûr* ». L'élève estime avec plus ou moins d'exactitude la qualité de ses réponses et l'état de ses connaissances, ce qui est adapté à l'évaluation des apprentissages. Il s'engage dans des activités d'autoévaluation de manière à développer sa confiance dans sa compréhension de ses processus d'apprentissage en vue d'une remédiation. Pour sa part, l'autoévaluation définie au sens strict est une démarche qui consiste à dire quelque chose de son travail en vue d'une remédiation (Wegmuller, 2002). Pour être valide dans un contexte de mesure

Tableau 2

*Taxonomie révisée (tiré et traduit de Anderson et al.; 2001)*

Dimension de la connaissance	Dimension du processus cognitif					
	Se souvenir (de mémoire)	Comprendre	Appliquer	Analyser	Évaluer	Créer
Connaissance factuelle						
Connaissance conceptuelle						
Connaissance procédurale						
Connaissance métacognitive						

et d'évaluation, il faut avoir confiance dans le jugement de l'élève et bien comprendre les facteurs qui peuvent influencer ou biaiser ce jugement. Quoique les activités d'autoévaluation soient plus souvent observées dans un contexte d'évaluation formative, ces mêmes activités d'autoévaluation sont tout aussi pertinentes dans des contextes diagnostiques et même certificatifs. Leclercq (1993, 1998) et Gilles (2005) font état de plusieurs exemples.

Leclercq (1993, 1998) et Gilles (2005), proposent plusieurs variations de la modalité de réponse permettant à l'étudiant d'exprimer sa certitude à l'égard de la qualité de ses réponses. La

Figure 2 en présente un exemple tiré de Boulé (2007). Notons que Leclercq et Gilles utilisent ces modalités auprès d'adultes dans un contexte universitaire; leur conception de ces modalités de réponse a évolué considérablement au fil des années. Certaines des modalités utilisées antérieurement utilisent une échelle de 0 à 100 % divisée par échelon de 20 points de pourcentage. Leclercq introduit une échelle qui discrimine davantage à ses extrémités comme ceci : 2 %, 10 %, 25 %, 50 %, 75 %, 90 % et 98 %. Parfois, ces échelles sont accompagnées de barèmes tarifaires, c'est-à-dire d'une pondération des items qui varie en fonction de la cohérence entre la qualité des réponses et la certitude exprimée. L'utilisation de tarifs est appliquée dans un contexte d'évaluation certificative. Gilles propose aussi un concept spectral des performances des étudiants. Leclercq soutient qu'une formation préalable à l'utilisation de ces échelles est requise. L'utilisation des modalités introduites par Leclercq et Gilles n'est pas nécessairement intuitive et n'est probablement pas adaptée à des élèves plus jeunes. À titre d'exemple, l'introduction des concepts de base qui portent sur les concepts liés aux probabilités et aux pourcentages n'est prescrite qu'à partir de la 6<sup>e</sup> année. La plupart des élèves de notre échantillon n'auront acquis les notions de proportion et de probabilité que bien après la sixième année, et ce malgré le fait qu'elles soient enseignées. Le chapitre qui porte sur la méthode introduit une alternative plus adaptée aux jeunes élèves. Retenons que l'évaluation du degré de certitude est complexe et elle doit être adaptée à la situation comme le niveau scolaire des élèves et la fonction de l'évaluation. Les barèmes tarifaires ne sont pas pertinents dans un contexte diagnostique.

Le degré de certitude permet d'accroître la différenciation entre les élèves lorsque le score composite combine la performance de l'élève à son degré de certitude exprimé. Il aide aussi à accroître la différenciation au niveau des apprentissages, car il facilite l'identification des items

<b>Évaluation diagnostique mathématiques</b>	
<p>Tu vas remarquer que, pour chaque question du test, il va falloir que tu dises si tu es certain de ta réponse. Ça ne change rien à ton résultat mais ça peut nous aider à mieux voir où tu es rendu. Il est important que tu répondes à toutes les questions et que tu indiques si tu es certain pour chacune des questions.</p>	
<p><b>6. Dispose les nombres ci-dessous en ordre croissant: 0,7 - 2/3 - 69%</b></p> <p><input type="radio"/> a) 0,7 - 2/3 - 69%</p> <p><input type="radio"/> b) 2/3 - 0,7 - 69%</p> <p><input type="radio"/> c) 2/3 - 69% - 0,7</p> <p><input type="radio"/> d) 0,7 - 69% - 2/3</p>	<p><b>Tu es certain(e) de ta réponse à :</b></p> <p><input type="radio"/> 2% - pas sûr du tout</p> <p><input type="radio"/> 10% - pas sûr</p> <p><input type="radio"/> 25% - un peu sûr</p> <p><input type="radio"/> 50% - ne sais pas</p> <p><input type="radio"/> 75% - plutôt sûr</p> <p><input type="radio"/> 90% - sûr</p> <p><input type="radio"/> 98% - complètement sûr</p>

Figure 2. Exemple de modalité d'expression du degré de certitude (tiré de Boulé, 2007).

qui sont répondus au hasard (Leclercq, 2006), augmentant ainsi l'intérêt diagnostique. De deux élèves ayant le même nombre de réponses exactes à un test, celui qui exprime des degrés de certitude qui sont davantage en relation avec la qualité de ses réponses devrait pouvoir se distinguer avec un score plus élevé. Leclercq (1993) et Leclercq, Boxus, Brogniez, Wuidar et Lambert (1993) recommandent fortement qu'un *feedback* soit présenté à l'étudiant sous forme de profil de réalisme. Nous constatons que Leclercq et Leclercq et al. postulent implicitement que ce *feedback* est formateur, même s'ils n'en font pas la démonstration de façon empirique. Notons que le *feedback* remis à l'étudiant vise à favoriser la régulation des apprentissages par l'élève. En évaluation diagnostique, on s'adresse prioritairement au personnel enseignant ou à des spécialistes pour faire d'autres types de régulation. Ceci distingue en partie l'évaluation diagnostique et l'évaluation formative. Un test diagnostique qui fournirait des résultats de

réalisme valides au personnel enseignant permettrait de suivre la progression des élèves, ce qui s'inscrit dans la régulation de l'enseignement.

Parallèlement à l'étude des degrés de certitude, il est intéressant de noter ce courant de recherche qui porte sur le « sentiment de savoir » (*Feeling of Knowing*). Les devis utilisent souvent le rappel dans un contexte de reconstitution de scène de crime (Allwood, Innes-Ker, Holmgren, et Fredin, 2008; Knutsson, Allwood et Johansson, 2011). Ces recherches tissent des liens entre la métacognition et la régulation des apprentissages (Keast, Brewer et Wells, 2007; Koriat, 1993; Koriat et Levy-Sadot, 2001; Maril, Simons, Mitchell, Schwartz et Schacter, 2003). Similairement, un courant en neurologie étudie le concept de « sentiment de savoir » et de prise de décision (Burton, 2007; Lehrer, 2009). Jans et Leclercq (1997) ainsi que Leclercq (2009) estiment cependant que les degrés de certitude se distinguent du « sentiment de savoir » qui semble ne s'appliquer que dans des conditions de rappel dans un contexte de questionnaires à choix multiple. Tout comme l'utilisation des degrés de certitude permet de concevoir de nouvelles formes d'évaluation diagnostique et formative, il est concevable que le « sentiment de savoir » puisse également être utilisé.

### **Le réalisme**

La certitude exprimée, dans un contexte d'évaluation diagnostique, est utilisée pour déterminer le réalisme de l'élève; le réalisme constitue une mesure *per se* qui a son utilité parce qu'elle donne des indications sur les connaissances partielles de l'élève, les capacités d'apprentissage de l'élève et ses capacités métacognitives. Le personnel enseignant a donc, avec des renseignements sur le réalisme, des outils de régulation lui permettant de modifier la cadence du cours ou de mettre en place des mesures de remédiation ciblées.

Il est clair qu'on s'éloigne de l'évaluation diagnostique à partir du moment où l'information recueillie sert à alimenter un *feedback* qui serait remis à l'élève à des fins de favoriser l'autorégulation du réalisme et des apprentissages. Cet aspect de l'évaluation formative dépasse largement le cadre de cette étude. Nous notons néanmoins plusieurs contributions sur le sujet (Allal, 1999, 2010; Butler, Karpicke & Roediger, 2008; Crooks, 1988; Delcomminette, Chapellier & Leclercq, 2012; Kulhavy, Yekovich & Dyer, 1976; Kulhavy, Stock, Hancock, Swindell & Hammrich, 1990; Laveault, 2004a, 2004 b, 2010; Leclercq, 1993, 2009; Leclercq, Denis, Jans, Poumay & Gilles, 1999; Noël & Leclercq, 2011; Nunziati, 1990).

Le réalisme des élèves peut être considéré à titre de trait métacognitif. Certaines analyses vérifieront la stabilité de ce trait d'un domaine cognitif à l'autre. Il est probable que ce trait est stable selon qu'on parle d'évaluation diagnostique ou d'évaluation formative. Cette recherche ne permet cependant pas de vérifier cette affirmation.

### **Normes de réalisme**

Leclercq (1993) recommande l'établissement de normes de réalisme de manière à pouvoir qualifier le réalisme et déterminer entre autres s'il est suffisant. Ces normes sont validées par la suite par Gilles (1996). Le Tableau 3 présente les normes établies à partir de vingt-huit examens administrés par la Faculté de psychologie et des sciences sociales de l'Université de Liège. Par exemple, un réalisme de 0,71 est considéré faible.

Il est clair que ces normes ont été établies dans un contexte universitaire avec des adultes formés à l'utilisation des degrés de certitude. Cette population est entièrement différente de la nôtre; pour cela, nous avons dû tenir compte d'un effet de sélection potentiel. Des normes locales

pour les élèves plus jeunes devront éventuellement être établies et validées afin de faciliter l'interprétation des résultats des recherches subséquentes.

Tableau 3

*Normes de réalisme en situation d'examen et d'auto-estimation (tiré de Gilles; 1996)*

L'INDICE DE RÉALISME EST

	s'il est		Performance	
IDÉAL	compris entre	0,96 et 1	accomplie par	1,6 % des étudiants
EXCELLENT		0,91 et 0,95		7,1 %
TRÈS BON		0,86 et 0,90		13,9 %
BON		0,81 et 0,85		18,2 %
MOYEN		0,76 et 0,80		18,0 %
FAIBLE		0,71 et 0,75		15,2 %
INSUFFISANT		0,66 et 0,70		10,7 %
TRÈS INSUFFISANT		0 et 0,65		14,7 %

### Difficulté des items

Les recherches effectuées sur des adultes (Boulé, 2007; Boulé et Laveault, 2011; De Finetti, 1965; Fabre, 1980; Gilles, 1996; Jacobs, 1974; Kleitman et Stankov, 2001; Leclercq, 1993; Stankov, 1998) rapportent que les items ou les tâches qui font appel à des niveaux supérieurs de la pensée ou qui sont difficiles entraînent de l'incertitude chez les étudiants. Howie et Roebers (2006) arrivent à des résultats semblables autant chez les enfants que chez les adultes.

L'étude du jugement métacognitif et du réalisme peut contribuer à jeter un éclairage nouveau sur les performances de différents groupes d'élèves. Par exemple, des élèves peuvent négliger de réviser leurs réponses ou prendre moins de temps pour remplir un questionnaire

d'examen parce qu'ils surestiment la qualité de leurs réponses; par contre, d'autres élèves peuvent s'acharner et perdre du temps sur certains items qu'ils maîtrisent parce qu'ils sous-estiment la qualité de leurs réponses. Ainsi, ces deux groupes peuvent obtenir des scores plus faibles pour des motifs différents.

Contrairement aux liens entre les concepts métacognitifs et la certitude qui font l'objet de plusieurs recherches chez les adultes (Delcomminette et al., 2012; Flavell, 1976; Leclercq, 1993, 2009; Leclercq et Gilles, 2001; Leclercq et Poumay, 2005; Noël et Leclercq, 2011), l'étude du lien entre la certitude et le développement cognitif des enfants n'en est qu'à ses débuts. Dans un contexte d'évaluation diagnostique, l'information de type cognitif et métacognitif peut informer le personnel enseignant et permettre de mieux orienter ou aiguiller l'élève vers des solutions de remédiation adaptées à ses besoins.

Fabre (1980) remarque que la certitude est liée aux connaissances des étudiants testés. Lorsque le niveau de performance est bas, l'étudiant peut estimer incorrectement la qualité de sa performance à partir d'indices non pertinents. Selon les données de Stankov (1998) et de Kleitman et Stankov (2001), plus le nombre de bonnes réponses de l'étudiant est élevé, plus celui-ci a confiance en ses choix. Kleitman et Stankov (2007) corroborent à nouveau ce constat et remarquent une corrélation entre la confiance en soi et le score au *Metacognitive Awareness Inventory* (Schraw et Dennison, 1994). Jacobs (1974) note que l'expression du degré de certitude peut varier en fonction de la difficulté des items. De Finetti (1965) souligne également que les items difficiles créent un degré d'incertitude. Ces études ne relèvent cependant pas que la difficulté des items et des tâches est relative au niveau de performance des étudiants. Il nous

apparaît que ce n'est pas autant le degré de difficulté de l'item ou de la tâche qui importe, mais l'écart entre le degré de difficulté et la capacité d'assimilation des étudiants.

Notons que les recherches mentionnées ci-dessus ont été réalisées auprès d'étudiants universitaires adultes de premier cycle. Dans l'optique de la métacognition, l'élève plus âgé a probablement une plus grande capacité à s'autoévaluer avec exactitude et réalisme que l'élève plus jeune. On peut aussi s'attendre à ce que les étudiants universitaires aient, pour la majorité d'entre eux, atteint un stade de développement cognitif avancé, du moins certainement plus avancé que celui des élèves du niveau élémentaire. Si le réalisme des étudiants adultes et des élèves relève de facteurs similaires, les questions de recherche et le devis permettront de vérifier si le réalisme peut varier selon les situations comme celles du niveau de difficulté des items et du domaine de ceux-ci. Par contre, si le réalisme des étudiants adultes et des élèves relèvent de processus différents, comme ceux liés à la maturité développementale, les résultats obtenus permettront de documenter le développement du réalisme en fonction du développement cognitif.

### **Développement cognitif**

Leclercq (1993), Fabre (1980), Stankov (1998), Kleitman et Stankov (2001), Jacobs (1974), De Finetti (1965), Boulé (2007), Boulé et Laveault (2011), Howie et Roebbers (2006) et Gilles (1996) montrent que les items qui font appel à des exigences cognitives élevées de niveaux supérieurs de la pensée ou qui sont difficiles entraînent de l'incertitude. Il est donc possible de déduire que le stade de développement atteint peut influencer la certitude de l'élève (Laveault, 2003; 2009). Dans cette étude, le développement cognitif de l'élève est surtout important dans la mesure où il permet d'estimer jusqu'à quel point l'écart entre les capacités développementales atteintes par l'élève et la complexité de la tâche peut influencer le degré de certitude de l'élève en

sa performance. En d'autres mots, l'élève n'aurait-il pas davantage une propension à l'incertitude lorsque la complexité de la tâche, ou son exigence cognitive, serait supérieure à son stade cognitif atteint?

Plusieurs liens potentiels peuvent exister entre les variables suivantes : difficulté, capacité, connaissance et stade de développement. Pour les élèves, la complexité cognitive d'une tâche dépend de leur capacité d'assimilation selon leur niveau de développement. C'est pourquoi l'évaluation qui compte utiliser le degré de réalisme de l'élève à des fins diagnostiques doit prendre en considération le niveau de développement dans l'appréciation du réalisme de l'élève.

Il y a peu d'écrits sur la nature du lien entre le niveau cognitif atteint par l'élève et son réalisme. Une recherche antérieure (Boulé, 2007; Boulé & Laveault, 2011) utilise le score total à un test pour déterminer la capacité cognitive des élèves. Il s'avère que cette méthode pose des problèmes de dépendance entre le trait mesuré et la matière testée; par exemple, les étudiants moins performants dans une matière peuvent exceller dans une autre; ils peuvent alors être en mesure d'avouer leur ignorance et être réalistes quant à la qualité de leurs réponses en géographie, mais pas en histoire.

Nous faisons l'hypothèse que le niveau de développement cognitif atteint par l'élève, plutôt que ses résultats dans différentes matières scolaires, constitue un cadre de référence plus fiable et plus approprié pour étudier le lien entre le réalisme et les capacités de l'élève. Le niveau de développement cognitif nous permet d'estimer l'écart entre les capacités développementales atteintes par l'élève et le niveau d'exigence cognitive de la tâche. Dans cette recherche, nous optons pour une approche diagnostique critériée fondée sur les niveaux piagétien de développement cognitifs à titre de référentiel. Laveault et Grégoire (2004) font état que

l'approche critériée a pour fonction d'évaluer la maîtrise par les élèves de certaines caractéristiques tirées d'un référentiel. Dans cette recherche, les tests diagnostiques sont constitués de tâches hiérarchisées selon leur niveau d'exigence cognitive et permettent aussi de déterminer le niveau cognitif atteint par l'élève. L'échelle de développement cognitif de Piaget constitue le référentiel.

En ce qui concerne le développement cognitif et l'évaluation critériée de l'élève, notons que Piaget distingue deux grands domaines de développement : le domaine logico-mathématique qui porte sur des objets discrets, sériés ou dénombrés, réunis sous forme de classes (Legendre-Bergeron et Laveault, 1980) et le domaine spatio-temporel ou infralogique; ce dernier, selon Legendre-Bergeron et Laveault, est constitué de parties ou d'éléments d'objets continus tels que l'espace et le temps. Puisque cette recherche utilise les stades de développement piagétien comme référentiel et que ces derniers font la distinction entre deux domaines de développement, cette recherche isole donc cette variable. Les écrits rapportent peu de recherches qui mettent en relation les domaines de développement et le réalisme. Un des rares exemples est celui de Jonsson et Allwood (2003) qui se sont intéressés au réalisme des adolescents en séparant les domaines de « connaissance du vocabulaire » et d'« habiletés logiques/spatiales ». Ils ont constaté que le réalisme varie en fonction du domaine cognitif, mais pas nécessairement en fonction des différences individuelles telles le sexe et le style cognitif. Jonsson et Allwood constatent que la surestimation s'explique par plusieurs facteurs et que ces facteurs peuvent varier en fonction du domaine. Selon leurs résultats, les élèves tendent à moins surestimer la qualité de leurs réponses dans le domaine « connaissance du vocabulaire » que dans le domaine « habiletés logiques/spatiales ».

## L'âge de l'élève

Piaget (1974; 1977), en se basant sur le facteur de l'âge et de l'épistémologie génétique de l'élève, suggère que la prise de conscience et l'abstraction réfléchissante sont liées au processus qui permet à l'élève de tirer de l'information de ses propres actions. Cette conceptualisation est croissante et s'inscrit dans les schèmes opératoires formels qui apparaissent vers l'âge de 11-12 ans. À cet égard, Hansson, Rönnlund, Juslin, et Nilsson (2008) relèvent une interaction entre l'âge de l'élève, son niveau d'habileté cognitive et son réalisme. Allwood et al. (2008) ont constaté que, dans un contexte de rappel après avoir été témoins d'un crime, le réalisme est similaire chez les élèves de 9 à 11 ans et les adultes. Par contre, après plusieurs séances de rappel, les élèves de 9 à 11 ans surestiment la qualité des rappels; ce phénomène ne se produirait pas chez les adultes (Knutsson et al., 2011).

De plus, on peut se demander si le réalisme se développe différemment dans chaque domaine, c'est-à-dire logico-mathématique, en opposition au domaine spatio-temporel. L'élève peut-il être plus réaliste dans un domaine en supposant que les tâches des différents domaines évoquent des habiletés cognitives du même niveau? En fait, plusieurs hypothèses sont possibles, l'une, entre autres, voulant que le niveau cognitif atteint par l'élève ne soit pas le même dans chaque domaine; cela pourrait influencer l'écart entre l'exigence cognitive de la tâche et le niveau cognitif atteint de façon différente dans chaque domaine.

Ainsi, dans la mesure où le réalisme varie en fonction des exigences cognitives des tâches, celles, plus exigeantes au niveau cognitif que celui atteint par l'élève, peuvent entraîner une réduction du réalisme de l'élève. De la même manière, le réalisme de l'élève pourrait

augmenter dans le cas où le niveau cognitif invoqué par les tâches serait inférieur ou égal à celui atteint par l'élève.

Cependant, nous postulons que cette relation n'est pas linéaire. Dans les cas où les exigences cognitives de la tâche seraient nettement supérieures au niveau cognitif atteint par l'élève, le réalisme de l'élève pourrait augmenter puisque les élèves pourraient être certains, et avec raison, de ne pas connaître la réponse. Dans la prochaine section, la Figure 4 montre conceptuellement la courbe en question.

Les écrits nous permettent de distinguer l'évaluation diagnostique de l'évaluation formative en mettant l'accent sur leur fonction primaire. L'évaluation diagnostique vise surtout à informer le personnel enseignant et favoriser la régulation de l'enseignement. L'évaluation formative informe davantage les élèves.

L'élève qui répond à un item d'un test n'a pas toujours une connaissance complète de la réponse et peut néanmoins en avoir une connaissance partielle. Les items qui sont corrigés de façon dichotomique ne permettent pas directement d'observer les cas de connaissances partielles. La mesure de réalisme, obtenue en calculant la correspondance entre la certitude exprimée par l'élève et la qualité de ses réponses, permet au personnel enseignant de vérifier l'état des connaissances des élèves et de mettre sur pied des stratégies de remédiation au niveau de l'enseignement le cas échéant. Par exemple, le personnel enseignant pourrait reconsidérer la pertinence des tâches qui ont des exigences cognitives nettement plus élevées que la capacité d'assimilation des élèves. Pour ce faire, une taxonomie qui incorpore les concepts de la métacognition est requise.

La plupart des recherches et des exemples qui portent sur l'utilisation des degrés de certitude sont liés à un contexte d'évaluation d'étudiants adultes. Des modalités de réponses ainsi que des normes de réalisme ont été introduites dans ce contexte. Il reste à voir si ces modalités et ces normes s'appliquent avec des élèves de niveaux primaire, moyen et intermédiaire.

## CADRE CONCEPTUEL

L'étude du développement cognitif chez l'élève présente un grand intérêt en ce qui a trait à l'examen du degré de certitude et du réalisme à des fins d'évaluation diagnostique; en effet, le niveau d'exigence cognitive d'un problème et le niveau de développement d'un élève peuvent être évalués quantitativement. Par exemple, en utilisant les stades de développement cognitifs piagétiens comme critères, il est possible de pister et déterminer les stades atteints par les élèves; de la même manière, il est possible d'établir le niveau d'exigence cognitive d'une tâche à partir des schèmes opératoires invoqués par cette tâche.

Dans une perspective développementale piagétienne, la cognition réfère à la perception, l'apprentissage, la pensée et la mémoire qui se développent chez l'humain selon un cheminement séquentiel prédéterminé selon un ordre invariant. Piaget (1930) exprime cette séquence sous forme de stades et de sous-stades marqués par des moments culminants observables. Cette recherche utilise ce cheminement cognitif à titre de référentiel critérié pour établir le niveau cognitif atteint par l'élève. Le tableau 4 qui suit résume brièvement ces stades en questions.

Piaget (1977) suppose un mécanisme de réfléchissement qui est en quelque sorte une réflexion sur le développement de la connaissance; il observe que l'élève peut dégager les caractéristiques de ses actions et de ses processus cognitifs; ainsi, pour l'élève, il s'agit d'une prise de conscience. Notons qu'il y a plusieurs notions chez Piaget qui rendent compte du réalisme dans un contexte développemental. Selon Piaget, l'abstraction réfléchissante comporte deux aspects : d'une part, un mécanisme de réfléchissement, soit une projection sur un palier supérieur de ce qui est tiré du palier inférieur; d'autre part, une réflexion sur les caractéristiques de ses actions et de ses processus cognitifs. Le concept d'abstraction réfléchissante s'apparente

Tableau 4

*Tableau résumant les stades de développement cognitif piagétien*

Stade	Âge approximatif	Description
L'intelligence sensori-motrice	0 à 2 ans	L'intelligence est basée sur l'action du corps et le milieu. Ce stade est caractérisé par le tâtonnement. Ce stade est composé de six sous-stades
L'intelligence préopératoire	2 à 7 ans	C'est le stade de la pensée symbolique. L'enfant développe l'imitation, la représentation, la réalisation d'actes fictifs. Un objet peut devenir le substitut d'un autre objet.
Le stade des opérations concrètes	7 à 11 ans	La pensée de l'enfant se socialise et il peut prendre en compte l'avis des autres. L'enfant développe les concepts de la causalité, le classement, la réversibilité, la temporalité et le regroupement.
Le stade des opérations formelles	11 ans à adulte	C'est le stade des opérations logiques, abstraites et du raisonnement hypothético-déductif.

alors à celui du réalisme en ce sens qu'il comporte les éléments de régulation des apprentissages et de la prise de conscience par l'individu de ses propres processus cognitifs.

Le processus mental lié au réalisme réfère à la conscience qu'a l'élève de ses propres activités cognitives et de ses doutes ou incertitudes. Piaget (1974) suggère différents niveaux de prise de conscience de la démarche cognitive; il traite de la prise de conscience comme d'une conceptualisation; il réfère à la capacité de se représenter, de décrire les phénomènes ainsi que le déroulement d'une démarche. Ce processus de prise de conscience, qui n'est d'ailleurs pas omniprésent dans toutes les activités cognitives, rejoint le processus métacognitif de Flavell (1976). La capacité à évaluer sa propre connaissance représente un aspect important de la

connaissance de soi. Cette définition peut être élargie en présupant que le réalisme de l'élève est une forme de connaissance métacognitive. Ainsi, l'expression de la certitude à l'égard des réponses choisies est une forme d'expression de jugement d'autoévaluation fortement empreint des capacités métacognitives de l'élève. (Allal, 1999; Laveault, 1998; Leclercq et Gilles, 2001).

Le réalisme métacognitif se mesure en établissant la concordance entre la certitude exprimée et la qualité des réponses. Les équations permettant d'en faire les calculs sont expliquées en détail dans le chapitre portant sur la méthodologie. L'élève peut être réaliste ou non en jugeant sa performance; par exemple, un élève peut systématiquement indiquer un degré élevé de certitude par rapport à des réponses erronées; on peut alors considérer que cet élève se surestime. Le Tableau 5 montre les croisements des cas possibles. L'élève réaliste a tendance à anticiper une bonne réponse lorsque la réponse est en effet bonne (A). Il anticipe également une mauvaise réponse lorsque celle-ci est mauvaise (D). L'élève peu réaliste qui sous-estime la qualité de ses réponses (C) a tendance à avoir une faible certitude par rapport à ses bonnes réponses. Il faut aussi prendre en considération que ces élèves qui sous-estiment la qualité de leur réponse peuvent faire partie du groupe qui répond au hasard et qui l'admet en exprimant une certitude faible. Un élève peut aussi sous-estimer sa réponse même s'il n'a pas répondu au hasard. Il peut manquer de confiance en soi ou hésiter entre deux réponses plausibles. Plus ils obtiennent de bonnes réponses grâce à la chance, plus leur score de réalisme diminue. L'autre type d'élève peu réaliste est celui qui surestime généralement la qualité de ses réponses (B). Il exprime une grande certitude pour des réponses incorrectes. Cette catégorie d'élèves mérite une attention particulière et une intervention qui favorise la régulation des apprentissages. On parle

en quelque sorte d'évaluation formative. Cette intervention peut prendre la forme d'un enseignement correctif.

Tableau 5

*Lien entre la réponse à l'item et le degré de certitude exprimé (tiré de Boulé, 2007)*

Lien entre la réponse à l'item et le degré de certitude exprimé		Qualité de la réponse à l'item (0/1)	
		Bonne (1)	Mauvaise (0)
Qualité de la réponse anticipée (Certitude)	Élevée	<b>A</b> <b>Réalisme</b>	<b>B</b> <b>Surestimation</b> <b>(remédiation utile et nécessaire)</b>
	Faible	<b>C</b> <b>Sous-estimation</b>	<b>D</b> <b>Réalisme</b>

L'analyse des différents cas de réalisme inclut l'interaction avec d'autres variables liées à la nature de la tâche. En évaluation, l'interprétation des scores obtenus par les élèves peut tenir compte à la fois des exigences de la tâche en matière d'opérations cognitives ainsi que du niveau de développement opératoire atteint, soit le stade cognitif atteint par l'élève. En ce qui concerne l'élève, l'interprétation des scores obtenus peut varier en fonction de sa propension à surestimer ou à sous-estimer la qualité de ses réponses. Notons que les cas de surestimation, qui consistent à croire une réponse bonne alors qu'elle est fautive, constituent la catégorie de situations critiques

pour lesquelles une régulation est la plus utile et nécessaire; il s'agit de situations où l'élève n'a pas la bonne réponse et ne le sait pas.

La Figure 3 présente un algorithme de modélisation du processus d'interprétation des résultats quand il est question de surestimation. Dans un contexte diagnostique, nous vérifions d'abord s'il s'agit d'un cas de surestimation à l'échelle d'un seul élève ou de tout le groupe d'élèves; ensuite, nous examinons s'il y a un écart important entre les exigences cognitives de la tâche et le stade développemental atteint par l'élève ou le groupe d'élèves. Les résultats de l'évaluation diagnostique administrée à l'ensemble du groupe d'élèves détermineront le type de régulation qui se fera à ce moment-là (soit au niveau du réalisme des élèves ou au niveau de l'apprentissage) selon qu'il s'agit de surestimation du groupe d'élèves ou d'un élève en particulier, et ce, dans la mesure où les exigences cognitives de la tâche n'excèdent pas le niveau développemental des élèves. La première question est de savoir si l'élève a une propension à surestimer la qualité de ses réponses. Dans de telles circonstances et dans un contexte formatif, l'élève peut bénéficier de ce *feedback* qui favorise une remédiation éventuelle au niveau individuel.

Si, par contre, la surestimation par l'élève se produit alors qu'il est réaliste la plupart du temps, il faut alors vérifier si le groupe d'élèves est réaliste par rapport à la tâche et s'il y a un écart important entre les exigences cognitives de la tâche et le niveau développemental des élèves. Dans un tel cas où les exigences de la tâche sont de niveau égal ou inférieur au stade atteint, la régulation individuelle au niveau de l'apprentissage s'applique. Dans le cas où tout le groupe d'élèves manque de réalisme par rapport à une tâche, cela peut bien être attribuable à des exigences cognitives démesurées ou à des cibles d'apprentissage peu réalistes par rapport au

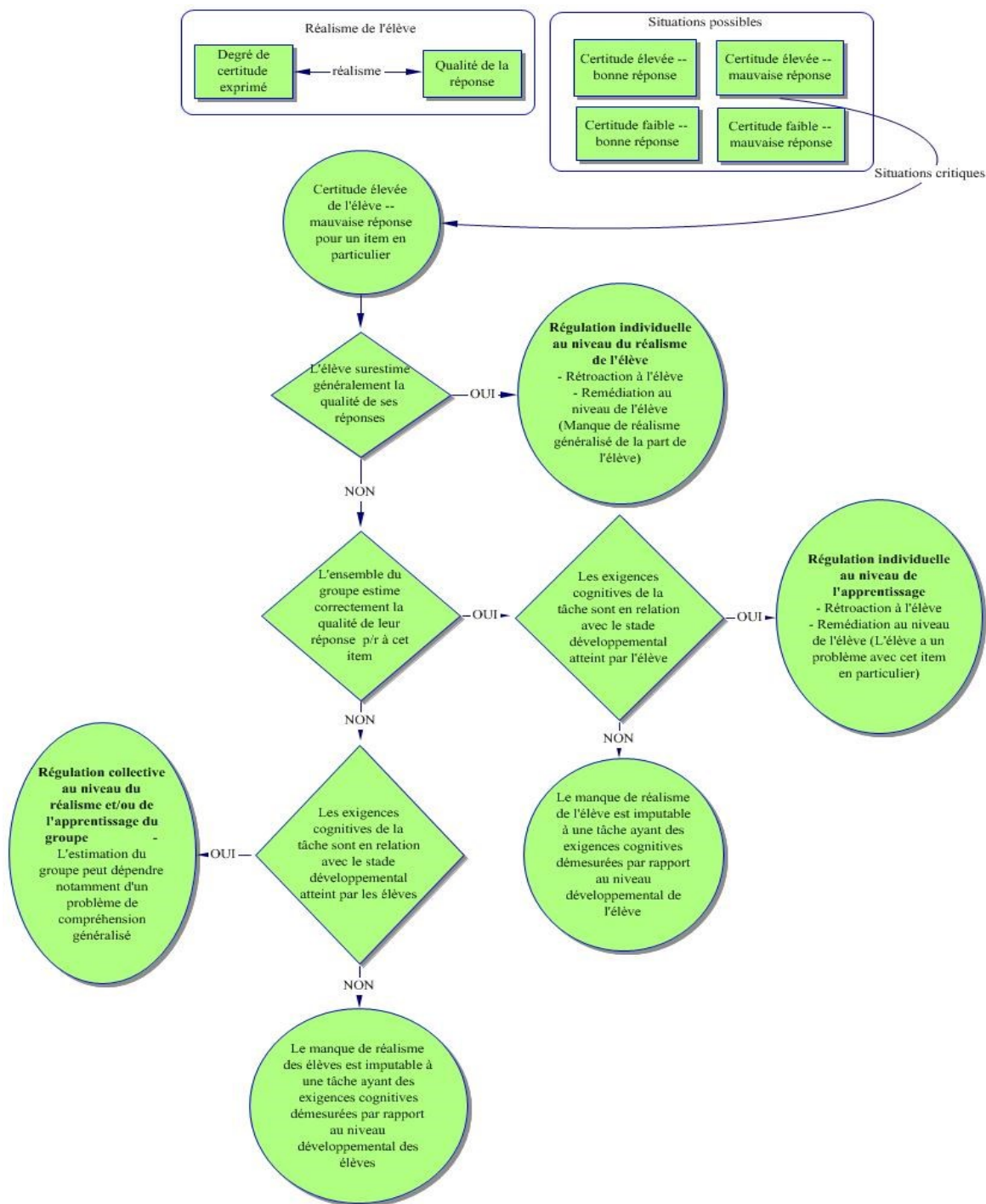
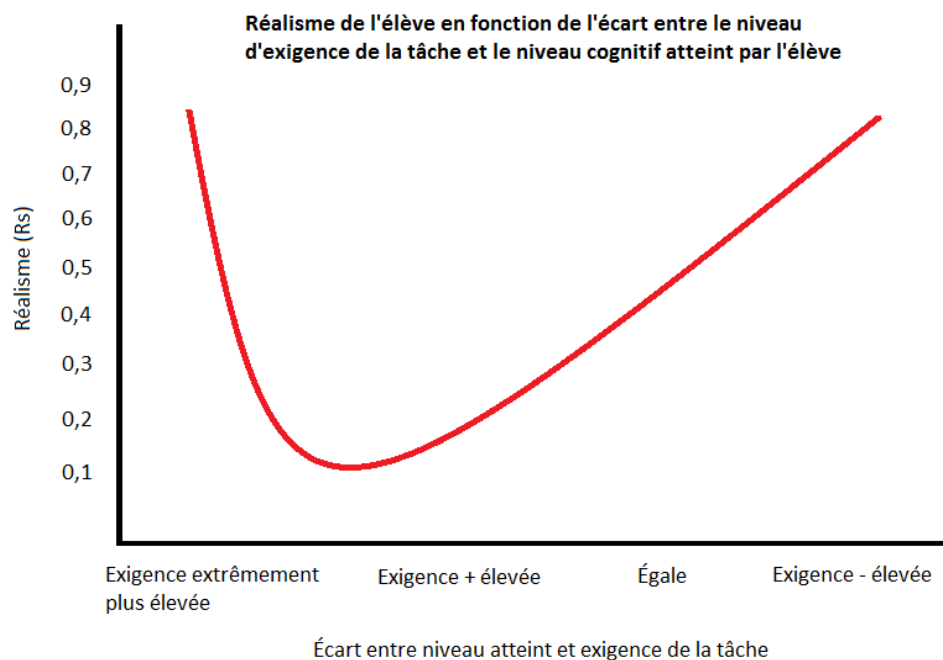


Figure 3. Algorithme de modélisation du processus d'interprétation des résultats en cas de surestimation.

stade développemental du groupe d'élèves. En d'autres mots, le manque de réalisme peut avoir plusieurs causes différentes. Conséquemment, la régulation des apprentissages des élèves peut être fort utile et s'inscrit dans le cadre du concept de zone proximale de développement (Vygotsky, 1986), car elle vise à offrir aux élèves une forme externe d'appui leur permettant de dépasser leur niveau de fonctionnement actuel dans la mesure où les cibles d'apprentissage sont réalistes. L'évaluation diagnostique n'est qu'une étape dans la collecte de renseignements permettant de meilleures régulations des apprentissages au niveau individuel et de groupe.

Voilà pourquoi nous estimons que l'élève est plus susceptible d'être réaliste si son niveau de développement cognitif est supérieur ou égal à l'exigence cognitive de la tâche. À l'inverse, l'élève risque d'être moins réaliste si son niveau de développement est inférieur à l'exigence cognitive de la tâche. Il est donc concevable et acceptable que les élèves montrent de l'incertitude devant une tâche trop exigeante par rapport à leur niveau de développement cognitif atteint. En d'autres mots, si l'exigence cognitive d'une tâche s'éloigne de la zone proximale (Vygotsky, 1986) de développement cognitif de l'élève, il est probable que l'élève éprouve de la difficulté à s'autoévaluer, et par conséquent, manque de réalisme. Cependant, cette relation n'est probablement pas linéaire. Dans les cas extrêmes d'items faciles et difficiles, l'élève pourrait manifester un haut degré de réalisme puisqu'il est certain de savoir ou de ne pas savoir, et ce, avec raison. On peut songer à une relation en forme de « U ». La Figure 4 qui suit montre conceptuellement l'allure possible de la courbe en question.

Boulé et Laveault (2011) font état de la complexité du jugement à porter sur le degré de certitude exprimé par l'élève et de la hiérarchie des difficultés liées à l'estimation de la réussite à une tâche ou aux capacités métacognitives de l'élève. Nous devons tenir compte de facteurs tels



*Figure 4.* Réalisme de l'élève en fonction de l'écart entre le niveau d'exigence de la tâche et le niveau cognitif atteint par l'élève.

que le sexe, le niveau d'habileté de l'élève, les propriétés métriques des items, tels le degré de difficulté et le coefficient de discrimination, et la nature de l'interaction entre ces facteurs qui peuvent influencer le réalisme de l'élève. Les écrits font état que le réalisme peut être affecté par des caractéristiques individuelles comme le sexe des individus. Pensons notamment aux travaux de Stankov et Lee (2008) et Stankov, Lee et Paek (2009) qui portent sur les différences de réalisme en fonction du sexe et de l'ethnicité. L'étude de certaines caractéristiques individuelles dépasse les objectifs de cette recherche. Certaines variables liées aux caractéristiques individuelles, comme le sexe de l'élève, sont néanmoins contrôlées statistiquement afin d'éviter que celles-ci ne confondent des effets expérimentaux.

Nous cherchons, entre autres, à vérifier si les exigences cognitives provenant de tâches des domaines logico-mathématique et spatio-temporel influencent le degré de certitude de la même façon. Il s'agit d'habiletés qui s'exercent dans des domaines différents et qui peuvent se manifester différemment selon les élèves.

De plus, notons que la majorité de ces recherches sur le réalisme a été effectuée dans un contexte universitaire; quelques-unes ont été réalisées avec des élèves des niveaux intermédiaire et secondaire (7<sup>e</sup> à la 12<sup>e</sup> année), le plus souvent avec des élèves des 11<sup>e</sup> et 12<sup>e</sup> années.

L'échantillon de participants des recherches sur le réalisme comprend surtout des élèves ou des étudiants qui, pour la plupart, ont atteint des stades de développement cognitif élevés. Ces recherches apportent peu ou pas de renseignements sur les élèves plus jeunes. En fait, nous savons peu de chose quant à quel âge ou à quel niveau cognitif atteint, l'élève peut autoévaluer *a posteriori* la qualité de ses performances avec réalisme, encore moins *a priori*. Par exemple, les recherches de Jonnaert et Laveault (1994) ont abordé la question de l'autoévaluation *a priori* de la familiarité d'une tâche; ceux-ci parviennent à la conclusion que les élèves d'âge moyen de 17 ans 6 mois peuvent le faire avec fidélité. Jonnaert et Laveault auraient-ils fait les mêmes constats avec des élèves de moins de 12 ans?

Le cadre conceptuel fait état que l'étude du réalisme peut grandement bénéficier du corpus de théories lié au développement cognitif de l'enfant. Nous pouvons d'abord établir le niveau de développement cognitif de l'élève. Nous pouvons ensuite déterminer l'exigence cognitive de la tâche. Nous pouvons donc prendre en considération l'éloignement de la tâche par rapport à la zone proximale de développement dans l'interprétation à accorder au réalisme de l'élève.

Le réalisme s'interprète en fonction du lien entre la certitude exprimée et la qualité de la réponse. L'insuffisance de réalisme peut être causée par la sous-estimation ou la surestimation de la qualité des réponses par l'élève. Nous notons également que le domaine de la tâche ainsi que le sexe de l'élève peuvent constituer des facteurs affectant le réalisme. L'intérêt diagnostique du réalisme et la régulation de l'enseignement doit prendre ces facteurs en considération.

## QUESTIONS DE RECHERCHE

Il est concevable que de l'incertitude chez les élèves soit engendrée par la difficulté d'un item faisant appel à des capacités cognitives éloignées de leur capacité d'assimilation; dans une telle situation, leur réalisme à l'égard de la tâche est-il affecté? La question fondamentale est la suivante : l'évaluation du réalisme a-t-il un intérêt diagnostique et peut-il donner lieu à des résultats valides, et si oui, dans quelles conditions de domaine cognitif et d'âge? Est-ce qu'il y a un fonctionnement différentiel en fonction des groupes d'appartenance des élèves? Voici les questions de recherche :

1. En contexte d'évaluation diagnostique, comment et dans quelle mesure le réalisme des élèves varie-t-il en fonction de l'écart entre la complexité de la tâche et leurs capacités cognitives selon leur stade de développement? Le réalisme des élèves change-t-il par rapport au domaine cognitif de la tâche, à savoir le domaine spatio-temporel ou le domaine logico-mathématique?
2. Dans les situations empreintes d'un manque de réalisme, la propension des élèves à surestimer ou à sous-estimer leurs capacités est-elle modérée ou influencée par leurs différences individuelles comme celles du sexe, du stade opératoire atteint ou du niveau scolaire?
3. Finalement, à partir de quel âge les élèves peuvent-ils autoestimer la qualité de leur performance avec réalisme? En ce qui concerne leur développement cognitif, où se situe leur zone proximale de réalisme? Dans une perspective d'épistémologie génétique, nous cherchons à déterminer l'âge minimal requis chez la majorité des élèves pour une

utilisation valide du réalisme dans un contexte scolaire. Il s'agit de déterminer le moment critique du développement cognitif des élèves où la majorité d'entre eux deviennent suffisamment réalistes pour considérer l'autoévaluation de la qualité de leurs performances comme une valeur ajoutée à la régulation de leur propre apprentissage.

Pour répondre à ces questions de recherche, nous soumettons les élèves à deux épreuves diagnostiques mesurant deux domaines cognitifs. Les élèves effectuent simultanément une tâche et sont interrogés sur leur degré de confiance dans l'exactitude de leurs résultats ou de leurs réponses (autoévaluation). Leur réalisme et leur niveau de développement cognitif peuvent être déterminés. Nous contrôlons le niveau d'exigence cognitive des tâches, le niveau scolaire et le sexe des élèves. La prochaine section présente la méthodologie employée.

## LA MÉTHODE

Pour répondre aux questions de cette recherche, nous avons choisi de faire appel à une base de données déjà existante. La base de données d'un Conseil scolaire, qui a comme secteur de fréquentation une région du nord de l'Ontario, possède des données pertinentes pour cette recherche. En plus de contenir des variables d'ordre cognitif et développemental, le contenu de cette base de données permet de mettre en relation l'âge de l'élève, son sexe et sa phase développementale atteinte. Cette base de données comporte les niveaux cognitifs atteints par les élèves dans les domaines spatio-temporel et logico-mathématique, les scores obtenus dans les deux épreuves et les degrés de certitude exprimés. Le contenu de cette base de données permet de construire de nouvelles variables, dont l'indice de réalisme qui constitue en quelque sorte une mesure des capacités métacognitives. Cette base de données comporte un échantillon de convenance stratifié constitué d'élèves des deuxième, cinquième et huitième années inscrits dans plusieurs écoles élémentaires; elle contient des données sur environ 40 % de la population totale du Conseil inscrite à ces niveaux scolaires.

Ce conseil scolaire s'est engagé dans une étude longitudinale en sollicitant notre collaboration. Le Conseil cherchait entre autres à se doter d'épreuves diagnostiques de dépistage. Le choix d'instruments s'est arrêté sur les instruments de Noelting (1980). Pour le Conseil, cette vaste étude longitudinale vise à analyser et identifier des prédicteurs du rendement des élèves dans une perspective de dépistage et d'amélioration du rendement des élèves. Le Conseil a procédé à l'administration des instruments. Cette recherche utilise donc des données secondaires.

Cette section documente la méthode et les mesures utilisées lors de la collecte des données par le Conseil. Elle présente d'abord les participants et les instruments de mesure utilisés. Elle discute ensuite de la procédure mise en place pour favoriser une cueillette rigoureuse des données. Finalement, elle explique les formules utilisées lors des calculs portant sur les variables de cette étude.

### **Les participants**

L'échantillonnage repose sur une approche développementale transversale. L'échantillon visé est constitué des cohortes d'élèves correspondant à trois groupes d'âge bien définis, à savoir des élèves de la deuxième année, de la cinquième année et de la huitième année. Les élèves de la deuxième année, âgés d'environ 7 ans, possèdent pour la majorité une capacité cognitive correspondant au niveau de pensée opératoire concrète inférieure (IIA); ceux de la cinquième année, âgés d'environ 11 ans, ont atteint en principe la période des opérations concrètes supérieures (IIB). Finalement, les élèves de la huitième année, âgés d'environ 15 ans, se situent au début du stade des opérations formelles (IIIA) ou sont en transition vers ce stade dans une proportion plus ou moins grande, soit de l'ordre de 33 % et moins, selon les recherches antérieures de Noeiting (1977).

Dans le cadre de cette recherche, nous utilisons une partie des données de la deuxième année de collecte par le Conseil scolaire; elles portent sur des élèves francophones qui ont le français comme langue maternelle et proviennent de milieux socio-économiques hétérogènes. Lors du tirage des données à partir de la base de données du Conseil, nous avons recueilli des données sur 679 élèves de manière à assurer des échantillons des différents stades d'environ 200

élèves (100 élèves \* 2 sexes \* 3 niveaux scolaires). Les données provenant des échantillons sont anonymes et ne permettent pas d'identifier les élèves, leur classe, le personnel enseignant ou leur école. Le nombre d'élèves permet une conformité aux exigences prescrites par les analyses statistiques utilisées. De plus, le nombre de filles et de garçons est sensiblement équivalent à l'intérieur de chacun des trois stades même si le nombre de filles excède quelque peu le nombre de garçons. Cette disproportion reflète la population de niveau scolaire réelle qui est constituée de 53,2 % de filles. Nous pouvons ainsi tenir compte de cette variable.

Le Tableau 6 ci-dessous présente quelques statistiques descriptives sur les échantillons. Le nombre total d'élèves pour l'épreuve des Figures graduées est différent de celui pour l'épreuve des Concentrations. Pour plusieurs analyses, nous n'avons retenu que les copies complètes. Les copies dont les réponses sont incomplètes à la modalité des degrés des certitudes (DC) ont été éliminées puisqu'elles sont intraitables. En fait, dix-neuf cas sur 679 ont été retirés de l'échantillon de l'épreuve des Figures graduées en raison de données manquantes et treize cas ont été retirés puisqu'ils apparaissaient aberrants en fonction de critères statistiques qui seront élaborés dans la section portant sur les résultats, ce qui a porté le total des cas analysés à 647. En ce qui concerne l'épreuve des Concentrations, cinquante-quatre cas au total ont été retirés sur 665 pour un total de cas analysés de 611. Les données manquantes n'étaient pas traitables puisque nous n'avions aucun point de référence pour effectuer des imputations.

### **Les instruments**

Les choix d'instruments ont été effectués à partir d'instruments existants et validés, soit l'épreuve des Figures graduées de Noeltling (1980) et l'épreuve des Concentrations de

Tableau 6

*Effectifs de l'échantillon*

Sexe	Année d'études			
	2 <sup>e</sup> année	5 <sup>e</sup> année	8 <sup>e</sup> année	N total
Filles — Épreuve des Figures graduées	124	105	119	348
Garçons — Épreuve des Figures graduées	109	91	99	299
Filles — Épreuve des Concentrations	111	103	115	329
Garçons — Épreuve des Concentrations	98	89	95	282

Noelting (1980). Dans les deux cas, il s'agit d'épreuves opératoires et hiérarchiques au sens des scalogrammes de Guttman (1941). Nous avons eu accès aux instruments de Noelting qui a accepté d'y apporter de légères modifications de manière à permettre la cueillette de données sur le degré de certitude des élèves à l'égard de leurs réponses choisies. Les modifications consistaient à ajouter aux épreuves de Noelting une modalité permettant à l'élève d'exprimer son degré de certitude selon un procédé qui convient aux élèves de ces groupes d'âge.

L'Appendice C présente la modalité d'expression du degré de certitude utilisée dans cette étude. Elle est plus adaptée aux élèves plus jeunes et il s'agit d'une modalité d'autoévaluation connue des élèves.

### Épreuve des Figures graduées

Noelting a développé l'épreuve des Figures graduées en adaptant les figures à reproduire utilisées par Piaget et Inhelder (1948) et en ajoutant des nouvelles. Il s'agit en quelque sorte de reproduire des figures de complexité croissante. Un barème et une grille analytique permettent la notation des traces laissées par l'élève. L'épreuve des Figures graduées de Noelting est conçue en fonction d'une échelle ordinale et hiérarchique portant sur la reproduction de figures géométriques qui varient progressivement en difficulté. Cette échelle se fonde sur des notions spatio-temporelles. Les épreuves de Noelting s'appuient sur des concepts forts notamment puisqu'ils sont appuyés par une théorie éprouvée ont été testés à maintes reprises. Ceci nous permet d'avoir confiance dans les données recueillies par le Conseil.

La grille d'analyse des traces des élèves tient compte des composantes structurales invariantes de la figure géométrique et des mécanismes fonctionnels permettant le passage d'un stade au suivant. Le Tableau 7 présente les tâches de reproduction de Noelting (1980) et leur lien avec les divers stades développementaux requis pour s'acquitter correctement de chacune des tâches.

Cette représentation associe un lien hiérarchique strict de la progression entre les stades développementaux et les exigences cognitives de cette tâche du domaine spatio-temporel. Il est également noté que l'élève pouvant s'acquitter d'une tâche d'un niveau d'exigence donné peut aussi exécuter toutes les tâches moins exigeantes sur le plan cognitif. Ce lien semble être corroboré par les résultats de l'analyse factorielle de Laveault (1984) effectuée sur les résultats de l'épreuve des Figures graduées.

Nous pouvons vérifier certaines propriétés métriques de l'épreuve des Figures graduées utilisées en recourant à l'analyse hiérarchique par scalogramme. Le but est de déterminer si les résultats des élèves donnent lieu à des scores reproductibles et si leur distribution ordinale respecte les caractéristiques d'une échelle de Guttman (Guttman, 1950). Il faut en outre vérifier l'unidimensionnalité de l'épreuve et nous assurer que l'échelle est cumulative. La section qui porte sur les résultats présente les statistiques obtenues.



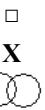


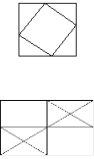
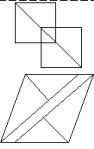
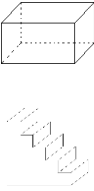
### Épreuve des Concentrations

L'épreuve des Concentrations porte sur les rapports de proportion entre différents éléments. Il s'agit d'une épreuve dans laquelle l'élève doit déterminer, par exemple, si le mélange constitué d'un verre de jus et de deux verres d'eau donne un mélange plus concentré que celui d'un verre de jus et de cinq verres d'eau. Cette tâche de ratio résolue par simple concentration sur la quantité de jus ou d'eau est considérée comme une tâche de niveau préopératoire. L'épreuve des Concentrations, au moyen de l'analyse des séquences, permet d'examiner le passage aux opérations formelles de la notion de proportion comme une succession hiérarchique de niveaux de fonctionnement (Laveault, 1983). En ce qui concerne cette épreuve des Concentrations, Laveault (1982) y trouve six regroupements factoriels correspondant aux stades développementaux. Ils sont présentés au Tableau 7.

En parallèle avec les Figures graduées, le Tableau 7 présente des tâches du domaine logico-mathématique telles que dans l'épreuve des Concentrations de Noelting (1980). À titre d'exemple, le regroupement IA correspond au stade intuitif supérieur. La figure à reproduire, celle du triangle, correspond à l'item (1,2) vs (1,5) du domaine logico-mathématique.

Tableau 7

*Découpage des stades par catégorisation et validation effectuée par analyse factorielle  
Tiré et adapté de Noeiting, Ferland, Laveault, Noeiting, Punde et Tessier (1984); Laveault (1982), Noeiting (1977) et Noeiting (1982)*

Regroupement par stade	Figure à reproduire	Nom de la figure	Spatio-temporel (Figures graduées)				Logico-mathématique (Concentrations)	
			Regroupement factoriel (Figures graduées)				Âge d'accession moyen (année, mois)	Items typiques (voir appendice)
			4	1	2	3		
S-M-sensori Moteur		Gribouillage Indifférencié						
OA-Symbolique Inférieur		Gribouillage Différencié Cercle					2,0	(1,0) vs (0,1)
OB-Symbolique Supérieur		Cercles disjoints Plus						
IA Intuitif Inférieur		Carré Multiplié Cercles sécants					3,6	(4,1) vs (1,4)
Moyen		Triangle					6,4	(1,2) vs (1,5)
IB-Intuitif Supérieur		Raquette Losange Cercles tangents					7,0	(3,4) vs (2,1)
IIA-Opératoire Concret Inférieur		Losange inscrit Drapeau					8,1	(1,1) vs (2,2)
IIB-Opératoire Concret supérieur		Carrés décalés Parallélogramme avec segments					10,5	(2,3) vs (4,6)
IIIA-Opératoire Formel Inférieur		Boîte en projection Oblique Escalier en projection oblique					12,2	(3,1) vs (5,2)
IIIB-Supérieur							14	(3,5) vs (5,8)

## **Domaines et stades développementaux**

Les épreuves de Noeiting (Noeiting, 1980; Noeiting et Gagné, 1980) s'avèrent intéressantes puisqu'elles peuvent être administrées collectivement plutôt qu'individuellement, ce qui permet de recueillir des données sur des échantillons de grande taille. Ces épreuves permettent aussi d'identifier les stades de développement piagétien à l'aide des concepts de raisonnement des domaines logico-mathématique et spatio-temporel. Les travaux de Noeiting (1982) et de Noeiting et al. (1984) ont permis de déterminer l'âge moyen associé aux différents stades tels que déterminés par les épreuves d'un domaine ou de l'autre. Afin de répondre à nos questions de recherche, les résultats à ces deux épreuves sont utilisés de manière à tester si le réalisme des élèves est indépendant du domaine. Quant aux différences individuelles, les différentes strates de l'échantillon transversal permettent de contrôler la phase développementale atteinte par les élèves de sept, dix et treize ans.

## **Propriétés métriques des instruments**

L'instrument utilisé par le Conseil scolaire est développé en fonction du modèle théorique de Noeiting et al. (1984) fondé sur les théories de Piaget. Il s'agit d'un modèle théorique fort qui assure la validité conceptuelle des instruments de mesure de la performance en fonction du développement des élèves. La validité théorique est documentée par plusieurs études (Bond, 1995; Laveault, 1981; Laveault, 1982; Laveault, 1983; Latouche-Garneau, 1977; Martel, 1976; Noeiting, 1982; Noeiting, Rousseau, Bond, et Brunel, 2000). Bond et Fox (2007) utilisent le modèle de Rasch pour identifier les stades de développement à l'aide de la théorie de la réponse aux items au lieu du scalogramme de Guttman et de l'analyse factorielle.

Dans l'épreuve des Concentrations, un coefficient de fidélité de 0,90 est obtenu en utilisant la formule de Spearman-Brown. Le coefficient de reproductibilité de Guttman (1950) se situe autour de 0,90 pour l'ensemble de l'épreuve (Noelting, 1982). Ces coefficients révèlent une scalabilité élevée et une structure reproductible des résultats, confirmant ainsi leur validité pour établir, au moyen de ces épreuves, les stades de développement atteints par les élèves. L'analyse factorielle exploratoire portant sur l'épreuve des Figures graduées révèle quatre facteurs distincts où les coefficients de saturation excèdent 0,40 (Laveault, 1984). Le coefficient de reproductibilité de Guttman se situe autour de 0,96 pour l'ensemble de l'épreuve (Noelting). Les items sont donc bien hiérarchisés.

En ce qui a trait à la validité conceptuelle, il s'agit de tâches conçues pour des élèves; leur utilisation est donc appropriée dans un contexte scolaire. De plus, une administration rigoureuse des épreuves ainsi qu'un processus de notation contrôlé ont contribué à obtenir des résultats fidèles et valides. L'administration de l'instrument, effectuée sur une base collective, a exigé environ 30 minutes.

### **La procédure**

Cette recherche a utilisé une approche par questionnaire où le participant fournit deux réponses à chaque question; la première réponse consiste à tenter de répondre correctement à l'item alors que la deuxième consiste à fournir le degré de certitude en la qualité de la réponse fournie. Les variables dépendantes ont toutes été enregistrées sur leurs échelles respectives durant la phase de *testing*. Il n'y a pas de conséquence irréversible ou d'intérêt à tricher. Le Conseil a effectué auprès d'un groupe d'élèves une mise à l'essai des épreuves et des modalités

de réponses de manière à s'assurer que les instruments ne présentent pas de vice de construction ou de reprographie.

Une échelle en cinq points de type Likert constituée d'images a permis à l'élève d'exprimer et de consigner ses choix. Les élèves utilisent régulièrement cette échelle et y sont habitués. L'échelle peut être consultée à l'Appendice C.

Les épreuves ont été administrées de façon collective dans les classes habituelles des élèves de plusieurs écoles d'une région urbaine du nord de l'Ontario sur une période de deux semaines au printemps 2012, par une même personne connue des élèves. Notons que même si ces épreuves sont utilisées individuellement la plupart du temps, nous trouvons des exemples d'administration collective comme chez Cloutier et Girard (1976) et des recommandations par Noelting (1982). Après l'obtention du consentement des parents par le Conseil, chaque élève était libre de ne pas participer aux épreuves même si ses parents y avaient consenti. La procédure utilisée par le Conseil lors de la collecte des données permet d'avoir confiance en la validité de celles-ci. Le montage de la base des données à partir d'un numéro d'identification différent assigné à tous les élèves permet de préserver leur anonymat ainsi que de connaître le sexe et l'année d'études de chaque élève.

Voici comment se sont déroulées les diverses étapes du protocole d'administration des épreuves.

- En utilisant du matériel concret sur l'épreuve des Figures graduées, l'élève a reçu une courte démonstration et s'est exercé sur quelques items d'essai de manière à s'assurer qu'il comprenait bien la tâche ainsi que les modalités de réponses.

- L'élève a procédé à la reproduction des Figures graduées.
- L'élève a indiqué sa certitude à l'égard de la qualité de sa reproduction.
- Ensuite, en utilisant du matériel concret sur l'épreuve des Concentrations, l'élève a reçu une courte formation et s'est exercé sur quelques items d'essai de manière à s'assurer qu'il comprenait bien la tâche ainsi que les modalités de réponses.
- L'élève a procédé avec l'épreuve des Concentrations.
- L'élève a indiqué sa certitude à l'égard de la qualité de ses réponses.
- La notation des copies pour l'épreuve des Concentrations s'est effectuée à l'aide d'un lecteur optique de marque.
- Trois juges qui utilisent les paramètres de notation de Noelting et al. (1984) ont effectué la notation des copies pour l'épreuve des Figures graduées. Toutes les copies ont été notées d'abord par deux juges, soit la psychométricienne du conseil et un assistant. Le troisième juge, qui est l'auteur de cette thèse, n'a procédé à une notation que lorsqu'il n'y avait pas une parfaite concordance des jugements des deux premiers juges. Le troisième juge devait donc prendre une décision dans le cas où les jugements des deux premiers juges étaient partagés.
- L'administrateur des systèmes de gestion de base de données du Conseil a consigné l'ensemble des données dans des chiffriers séparés par épreuve. Il a ensuite procédé à la fusion des deux chiffriers à l'aide du numéro d'identification des élèves de manière à

produire les séries de données utilisées dans cette étude. Ces séries de données ne comportaient pas de données personnelles permettant d'identifier les élèves.

### **Le plan d'analyse des résultats**

En plus de ces concepts d'ordre métacognitif, nous avons mis en relation la phase développementale atteinte par l'élève; son niveau cognitif dans deux domaines est mis en rapport avec son réalisme.

Le sommaire des variables présenté au Tableau 8 permet de distinguer les variables selon leurs caractéristiques et leur type. Dans les diverses colonnes, les variables sont organisées en fonction de leurs caractéristiques. Il peut s'agir des caractéristiques individuelles des participants comme celles du sexe ou du niveau développemental atteint par l'élève. Les caractéristiques de la tâche incluent le niveau d'exigence cognitive des items. Les caractéristiques de la réponse réfèrent à l'exactitude des réponses, à la certitude exprimée envers ces dernières et au réalisme calculé des élèves ou des items.

Une telle classification permet de mieux répartir chacune des variables; celles-ci sont également ordonnées selon le type d'échelle qui peut être catégorielle ou continue. Les catégories possibles des variables catégorielles sont spécifiées. Les deux variables construites sont de type continu. Il s'agit d'abord du réalisme des élèves obtenu en établissant une relation entre leur certitude exprimée et la qualité de leurs réponses. Nous calculons ensuite le réalisme des élèves par rapport aux items, ce que Gilles (2002) appelle *Indice de réalisation des prédictions de l'item (Rq)*.

Tableau 8

*Sommaire des variables*

Caractéristiques	Types des variables		
	individuelles	de la tâche	de la réponse
Nominale	— Sexe	— Domaine cognitif (spatio-temporel; logico-mathématique)	— Exactitude de la réponse (1; 0)
Construite/ Continue ou ordinaire	— Niveaux développementaux atteints par l'élève (obtenus dans les deux épreuves) — Année d'études (2 <sup>e</sup> , 5 <sup>e</sup> , 8 <sup>e</sup> années)	— Indice de réalisation des prédictions de l'item ( $Rq$ ) ( $\geq 0$ et $\leq 1$ ) — Niveau d'exigence cognitive	— Réalisme des élèves ( $R_s$ ) ( $\geq 0$ et $\leq 1$ ) — Réalisme des élèves ( $R_s$ ) selon le niveau cognitif évoqué des items ( $\geq 0$ et $\leq 1$ )
Continue	— Âge		— Degré de certitude (DC) (variable continue catégorisée selon les probabilités = 0,02; 0,24; 0,50; 0,74; 0,98)

Une échelle nominale distingue le sexe des élèves ainsi que le domaine de la tâche qui peut être spatio-temporelle ou encore logico-mathématique. Des échelles ordinales classent les

données de l'année d'études des élèves, de leur stade de développement cognitif atteint selon leurs scores aux épreuves de Noelting (1980) et du niveau d'exigence cognitive de la tâche exprimée sous forme de stades de développement cognitif évoqués. Une échelle continue présente l'âge des élèves et leur degré de certitude exprimé par rapport à leurs réponses. Le réalisme des élèves, une variable construite de cette étude, apparaît sur une échelle de rapport.

De manière à répondre aux questions de recherche et à faciliter les analyses, les données sont disposées sous forme de deux matrices. La première matrice de données est organisée en fonction des réponses individuelles (S pour sujets ou élèves), du niveau de performance des élèves et des calculs liés au réalisme des élèves. Les élèves sont aussi catégorisés selon leur stade de développement cognitif atteint et leur niveau scolaire. Le sexe des élèves est considéré à titre de variable attribut. Depuis plusieurs années, les résultats aux tests provinciaux administrés par l'Office de la qualité et de la responsabilité en éducation (OQRE, 1999) font état de performance différentielle entre les garçons et les filles. Cette recherche ne tente pas de l'expliquer. Néanmoins, nous devons contrôler statistiquement cette variable de manière à empêcher qu'elle ne brouille, supprime ou dissimule d'autres relations entre les variables.

Quant à la deuxième matrice de données, elle met en évidence les demandes cognitives des tâches exprimées sous forme de niveaux développementaux évoqués (Q pour questions) ainsi que le réalisme associé à chaque tâche ( $Rq$ ) et à chaque domaine. À ce sujet, Gilles (2002) réfère à l'indice de réalisation des prédictions plutôt qu'au réalisme des questions en expliquant que les items d'une épreuve ne peuvent pas faire preuve de réalisme ou de métacognition au même titre que peut le faire un élève.

### Calcul du réalisme et de l'indice de réalisation des prédictions

Leclercq (1993) estime que le calcul de l'indice de réalisme est basé sur la concordance entre l'exactitude des réponses de l'élève et son degré de certitude exprimé. Plus il y a présence de concordance de ces deux variables, plus le réalisme de l'élève est élevé. Chaque fois qu'il y a absence de concordance, l'indice de réalisme diminue un peu. Cette diminution se chiffre à partir du calcul de l'erreur moyenne absolue de certitude ( $\bar{e}$ ) déterminée au moyen du calcul de la différence entre le taux d'exactitude ( $T$ ) et la valeur de la certitude ( $C$ ). Cette erreur doit être calculée pour chaque degré de certitude et sa moyenne représente le « non réalisme ». Cette recherche utilise cette formule pour calculer le réalisme des élèves ( $R_s$ ) et l'indice de réalisation des prédictions des items ( $R_q$ ). L'Appendice D fournit la formule pour le calcul de l'indice de réalisation des prédictions par question ( $R_q$ ).

Des équations dérivées de ce calcul ont été utilisées dans cette recherche pour calculer le réalisme des élèves ( $R_s$ ) et l'indice de réalisation des prédictions des stades développementaux ( $R_q$ ), ou en d'autres mots, le réalisme associé à une tâche. Leclercq (1993) ainsi que Gilles (1996, 2002, 2005) expliquent de manière détaillée les formules à utiliser, les variables à considérer ainsi que les postulats qui sous-tendent leur utilisation.

$$\bar{e} = \frac{\sum_{i=1}^n (|C_i - T_i| \cdot U_i)}{J_i}$$

$\bar{e}$  = erreur moyenne absolue de certitude

$i$  = indice des degrés de certitude

$n$  = nombre des degrés de certitude

$C_i$  = valeur de la certitude  $i$  (l'échelle doit permettre de l'exprimer en pourcentages)

$R_i$  = nombre de réponses correctes pour la certitude  $i$

$U_i$  = nombre d'utilisations de la certitude  $i$

$T_i$  = taux d'exactitude de la certitude  $i$  (en pourcentages) =  $(R_i / U_i) \times 100$

$J$  = nombre total de réponses =  $\sum_i U_i$

Et l'indice de réalisme ( $R_s$ ) sur une échelle de 0 % à 100 % correspond à :

$$RÉALISME(R_s) = 100 - (\bar{e})$$

## RÉSULTATS

Le plan d'observation, quoique complexe, ne comportait que deux variables dépendantes construites à partir de plusieurs autres variables dépendantes, et ce, pour chacune des épreuves. Les deux variables dépendantes construites, soit l'indice de réalisme des élèves ( $R_s$ ) et l'indice de réalisation des prédictions de l'item ( $R_q$ ), ont été traitées indépendamment selon la matrice analysée (soit R ou Q). La première matrice (Q) présente les données en fonction des items et met l'accent sur l'indice de réalisation des prédictions de l'item ( $R_q$ ). La deuxième matrice (R) fait ressortir le réalisme des élèves ( $R_s$ ); elle est organisée en fonction des élèves avec l'accent mis sur les variables individuelles qui leur sont associées ainsi que sur le domaine cognitif évoqué par les items. L'indice de réalisme ( $R_s$ ) et l'indice de réalisation des prédictions de l'item ( $R_q$ ) sont des scores composites calculés à partir du degré de certitude exprimé et la qualité de la réponse. Les équations sont présentées dans la section portant sur la procédure.

Nous avons retenu, à des fins exploratoires, un seuil de signification de  $\alpha = 0,05$  pour une erreur de type I. La taille des effets expérimentaux ( $\eta^2$  carré) sera utilisée pour déterminer quelles variables significatives seront prises en considération. Avec de grands échantillons, de faibles effets expérimentaux s'avèrent significatifs à cause de la trop grande puissance du test. Notons que l'erreur de type I (5%) fait en sorte que 5% des tests de signification conduiront à des conclusions erronées.

Les mesures obtenues à partir de la base de données fournie par le conseil scolaire ont été soumises à plusieurs analyses statistiques. Voici le sommaire des analyses qui ont été effectuées :

1. L'analyse de la fidélité inter-juges a permis d'établir et de vérifier la fidélité de la notation des items de l'épreuve des Figures graduées.
2. Les distributions de données ont été vérifiées de manière à repérer les données manquantes et les cas aberrants.
3. Les coefficients de reproductibilité ( $CR_i$ ) ont été calculés pour vérifier la reproductibilité des scores des épreuves des deux domaines, Figures graduées et Concentrations.
4. Les distributions de l'ensemble des variables indépendantes et dépendantes ont été examinées et commentées.
5. Les questions de recherche ont été vérifiées au moyen d'une série d'analyses de variance multivariées et univariées, ainsi qu'au moyen d'analyses de régression multiple. Plusieurs analyses ont été effectuées afin de comprendre les effets potentiels sur le réalisme dus au sexe et au stade développemental atteint par l'élève ainsi qu'à la complexité et au domaine cognitif de la tâche. La conformité des distributions aux postulats des analyses statistiques utilisées est vérifiée au préalable.

### **La fidélité interjuge**

La notation des copies pour l'épreuve des Figures graduées a été exécutée par 3 juges formés à l'utilisation des paramètres de notation de Noelting et al. (1984). En plus de l'auteur de cette thèse, les deux autres juges connaissaient les épreuves de Noelting et travaillaient dans le domaine de l'enfance en difficulté dans le conseil d'où provient l'échantillon. Notons que la personne ayant administré et noté les épreuves est psychométricienne et a été formée pour administrer les deux épreuves et noter l'épreuve des Figures graduées. Les copies d'ancrage

provenant des paramètres de notation de Noelting ont servi à la formation des juges et à titre de référence durant le processus de notation. Ces copies d'ancrage sont des exemples de figures reproduites qui servent à illustrer les cas limites d'acceptation ou de rejet.

Chaque figure reproduite a été notée sur une échelle dichotomique, soit 0 lorsque la reproduction était rejetée et 1 lorsque celle-ci était retenue. Toutes les copies ont d'abord été notées par deux juges seulement. Dans l'ensemble, 1,42 % des jugements des deux premiers juges ne concordaient pas parfaitement. Le troisième juge a évalué seulement les copies pour lesquelles il y avait absence de concordance parfaite des jugements des deux premiers juges; dans une telle éventualité, son jugement servait à établir une majorité et à départager la décision.

### **Les cas aberrants**

De manière à éviter certains biais dans les résultats des analyses, une attention particulière doit être portée envers les cas extrêmes et aberrants. Après avoir examiné les distributions des variables ayant des données de type continu, soit le réalisme des élèves, nous identifions quelques cas de données aberrantes. Ces cas sont également analysés dans une perspective multivariée puisque nous utilisons entre autres des méthodes d'analyse multivariée de la variance. Nous déterminons ensuite si nous devons éliminer ces cas. Les résultats de l'analyse de la distance de Mahalanobis des cas aberrants sont présentés ultérieurement.

Les Figures 5 et 6 font d'abord état des cas aberrants de type univarié. Ces figures présentent le réalisme des élèves pour les deux épreuves en fonction du niveau scolaire et du sexe de l'élève. Dans le cas de l'épreuve des Figures graduées, nous notons cinq cas aberrants chez les filles de 2<sup>e</sup> et 5<sup>e</sup> année et deux cas chez les garçons de 5<sup>e</sup> et 8<sup>e</sup> année. Il n'y a que deux cas chez

les garçons pour l'épreuve des Concentrations. Nous remarquons qu'il semble y avoir un effet de plafonnement de la performance surtout chez les élèves de 8<sup>e</sup> année. Ce phénomène est discuté ultérieurement dans la section sur les résultats.

L'échelle des niveaux cognitifs atteints est ordinale sans être proportionnelle. Puisque nous ne pouvons pas traiter cette échelle à titre de métrique comme dans le cas des échelles continues, nous avons éliminé par observation quelques cas qui apparaissaient extrêmes. Il semble y avoir plus de cas de données aberrantes chez les garçons que chez les filles. Nous remarquons aussi que plusieurs élèves se trouvent aux stades sensorimoteur ou symbolique, ce qui est étonnant considérant le niveau scolaire des élèves. Le conseil scolaire a confirmé que tous les élèves ont participé aux épreuves. On nous a indiqué que certains élèves ont une motricité réduite, un trouble d'attention ou même un retard de développement.

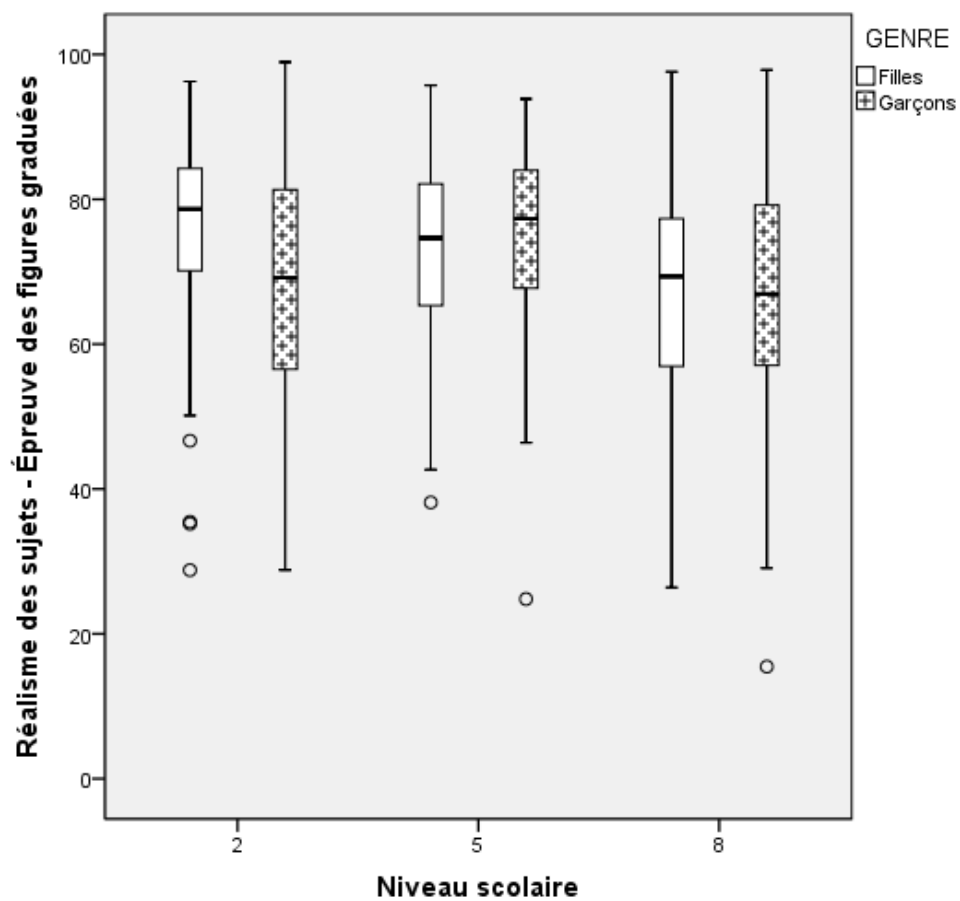


Figure 5. Distribution du réalisme des élèves pour l'épreuve des Figures graduées.

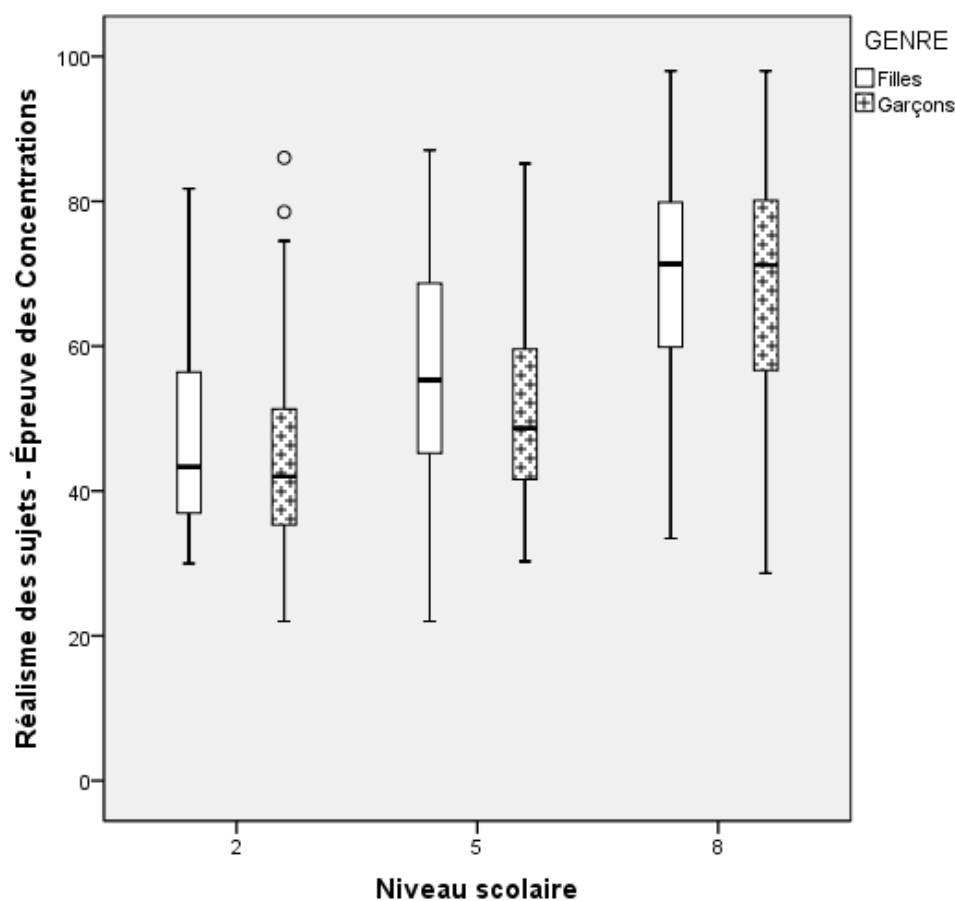


Figure 6. Distribution du réalisme des élèves pour l'épreuve des Concentrations.

Pour les variables sur des échelles continues, le choix d'exclure des cas aberrants se fait en conjonction avec les résultats de l'analyse de la distance de Mahalanobis des cas aberrants dans une perspective multivariée. Cette analyse est présentée ultérieurement. L'échelle des niveaux cognitifs atteints utilise un système de codes. Ces codes sont présentés au tableau 7.

Une analyse de la distance de Mahalanobis a permis de détecter douze cas qui s'avèrent aberrants sur le plan multivarié. Ces cas, d'ailleurs surreprésentés dans les stades des opérations formelles, ont donc été retirés. Il s'agit probablement des cas faux positifs dans les résultats de

l'épreuve des Concentrations. À cet égard, Noelting (1980a, 1980b) suggérait qu'on demande une justification de la part de l'élève par rapport à ses réponses pour les items de type IIIA et IIIB. En ce qui concerne cette recherche, il n'a pas été possible pour des raisons de faisabilité de se conformer à cette recommandation. L'analyse de distance de Mahalanobis permet néanmoins de détecter et d'exclure une partie de ces cas. Quelques cas aberrants ont été repérés et éliminés puisqu'ils se situaient à des stades de développement très inférieurs dans l'épreuve des Concentrations par rapport à l'épreuve des Figures graduées. Ces cas peuvent nuire au moment des analyses et des tests d'hypothèses. Notons néanmoins que le conseil scolaire est particulièrement intéressé à ces cas aberrants et considère que les deux épreuves sont très utiles dans une perspective diagnostique pour les élèves ayant des besoins particuliers.

### **Les coefficients de reproductibilité**

Le Tableau 9 affiche des coefficients de reproductibilité ( $CR_i$ ) pour les stades évoqués qui sont élevés pour l'épreuve des Figures graduées, avec un seuil de 0,90 comme étant souhaitable. Nous notons à l'aide du calcul du coefficient de reproductibilité marginal par stade évoqué ( $MMR_i$ ) que les items du stade des opérations formelles sont très difficiles et qu'ils gonflent artificiellement le  $CR$  à 0,912 pour l'ensemble de l'épreuve. Par ailleurs, l'indice de scalabilité ( $PPR$ ) est de 0,795, ce qui excède le seuil acceptable de 0,70 selon Jackson (1949). L'épreuve utilisée vise à donner une chance aux élèves de montrer ce qu'ils peuvent faire. Il est attendu que les items évoquant le stade III soient largement échoués. Nous savons que ces épreuves sont validées et qu'elles fournissent les mesures nécessaires aux fins de cette recherche. Notons que les coefficients de reproductibilité pour les deux épreuves sont légèrement plus élevés que ceux que Noelting (1982) a observés avec l'épreuve des Concentrations.

Le Tableau 10 présente les coefficients de reproductibilité pertinents pour l'épreuve des Concentrations. Les conclusions tirées pour l'épreuve des Concentrations sont semblables à celles de l'épreuve des Figures graduées.

Tableau 9

*Coefficients de reproductibilité pour les 16 items de l'épreuve des Figures graduées*

Paramètres	Stades évoqués						
	OA	OB	IA	IB	IIA	IIB	IIIA
$CR_i$	1,000	0,997	0,965	0,907	0,788	0,862	0,869
$MMR_i$	0,934	0,837	0,779	0,561	0,523	0,250	0,131
$PPR_i$	1,000	0,981	0,84	0,789	0,556	0,815	0,849

$$CR = 0,912$$

$$PPR = 0,795$$

Tableau 10

*Coefficients de reproductibilité pour les 15 items de l'épreuve des Concentrations*

Paramètres	Stades évoqués							
	IA	IB	IC	IIA	IIB	IIIA	IIIA2	IIIB
$CR_i$	0,991	0,962	0,943	0,971	0,995	0,992	0,997	0,995
$MMR_i$	0,974	0,749	0,519	0,299	0,105	0,083	0,045	0,026
$PPR_i$	0,647	0,850	0,881	0,959	0,995	0,992	0,997	0,995
$CR = 0,980$								
$PPR = 0,970$								

### **Les distributions de l'ensemble des variables indépendantes et dépendantes**

Avant de procéder à la vérification de nos hypothèses, on examine les distributions des variables de manière à bien comprendre comment elles se distribuent. La conformité aux postulats sera analysée dans la prochaine section.

#### Indice de réalisation des prédictions ( $Rq$ )

Les Figures 7 et 8 ci-dessous présentent l'indice de réalisation des prédictions pour chaque item des deux épreuves. Rappelons que l'indice de réalisation des prédictions représente le réalisme des élèves associé à un item en particulier. Par exemple, quel niveau de réalisme des élèves évoque cet item? Les items sont ordonnés de façon croissante en fonction de leur niveau

d'exigence cognitive. Le Tableau 7 dans le chapitre précédent montre le type de figures graduées et les différents scénarios de l'épreuve des concentrations.

Nous notons qu'il y a beaucoup plus de variations du  $Rq$  pour les items de l'épreuve des Concentrations surtout pour les items plus exigeants. Les items pr9 à pr14 sont les items des niveaux IIIA et IIIB pour lesquels nous n'avons pas pu obtenir de justification de la part des élèves. Notons également que l'épreuve des Concentrations comporte en moyenne plus d'items qui font appel à des niveaux de développement cognitif de type IIIA et IIIB. Il est attendu que les élèves de 2<sup>e</sup> année et de 5<sup>e</sup> année ont plus de difficulté à autoestimer la qualité de leurs réponses rendus à de tels niveaux d'exigence cognitive.

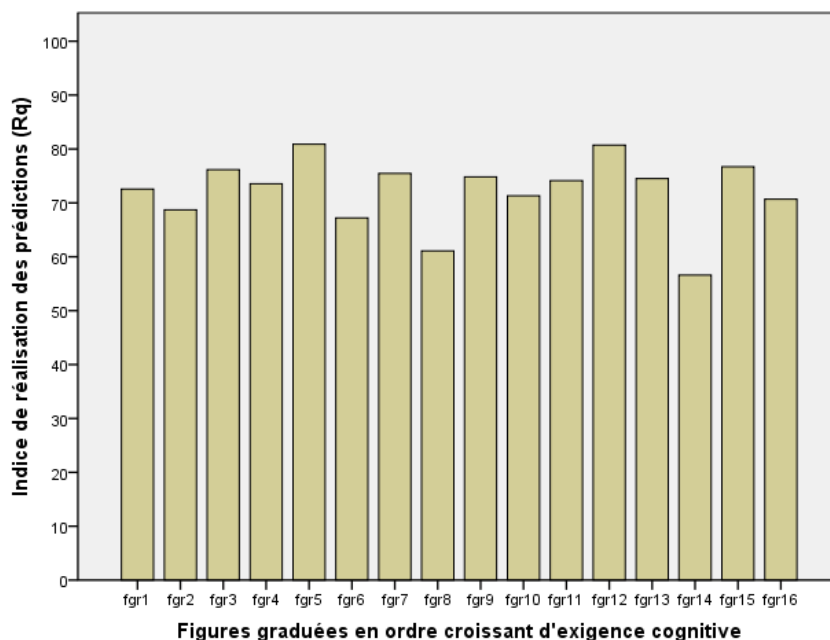


Figure 7. Indice de réalisation des prédictions de l'item ( $Rq$ ) pour l'épreuve des Figures graduées.

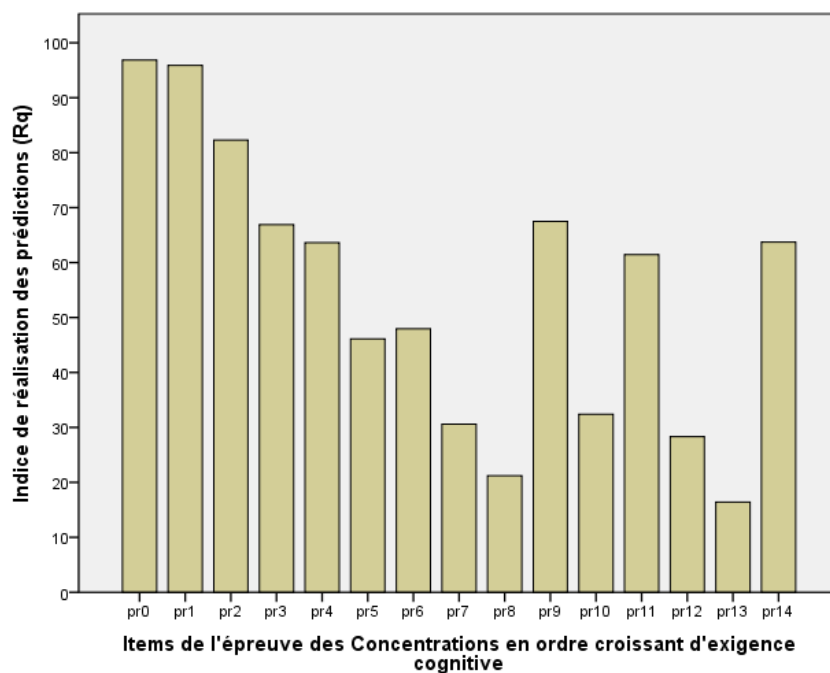


Figure 8. Indice de réalisation des prédictions de l'item ( $R_q$ ) pour l'épreuve des Concentrations.

### Stades cognitifs atteints

La Figure 9 montre la distribution des stades cognitifs atteints selon l'épreuve des Figures graduées pour l'ensemble des élèves, soit les élèves de la 2<sup>e</sup>, 5<sup>e</sup>, et 8<sup>e</sup> année. À première vue, nous constatons une légère asymétrie négative. Le Tableau 11 suivant fait état des coefficients d'asymétrie et de kurtose pour les élèves de chaque niveau scolaire tout en séparant les garçons et les filles. Nous constatons que les coefficients d'asymétrie sont tous négatifs et que la moitié d'entre eux sont significativement négatifs (selon le critère  $z = 2,58$ ). Ces observations suggèrent qu'il y a eu plafonnement de la performance surtout chez les élèves de 8<sup>e</sup> année. Nous notons également avec la Figure 9 que le nombre d'élèves est élevé au stade IIA. Pour le stade IIA, le

$CRi$  est égal à 0,788 ce qui est faible et inférieur au seuil de 0,85. Ceci suggère que le stade IIA pourrait être un stade de transition pour cette population d'élèves.

La Figure 10 montre la distribution des stades cognitifs atteints selon l'épreuve des Concentrations pour l'ensemble des élèves, soit les élèves de la 2<sup>e</sup>, 5<sup>e</sup>, et 8<sup>e</sup> année. Nous constatons une légère asymétrie positive. Le Tableau 12 fait état des coefficients d'asymétrie et de kurtose pour les élèves de chaque niveau scolaire tout en séparant les garçons et les filles. Nous constatons que les coefficients d'asymétrie sont presque tous positifs sans pour autant être significatifs. La distribution pour les filles de 2<sup>e</sup> année est significativement platykurtique. Ceci suggère que ce groupe d'élèves est très hétérogène et que plusieurs élèves ont mieux performé que ce qui était attendu.

#### Réalisme des élèves

La Figure 11 montre la distribution du réalisme des élèves pour l'épreuve des Figures graduées, et ce, pour l'ensemble des élèves, soit les élèves de la 2<sup>e</sup>, 5<sup>e</sup>, et 8<sup>e</sup> année. Nous constatons une légère asymétrie négative. Le Tableau 13 suivant fait état des coefficients d'asymétrie et de kurtose pour les élèves de chaque niveau scolaire tout en séparant les garçons et les filles. Nous constatons que les coefficients d'asymétrie sont tous négatifs et significatifs dans le cas particulier des filles de la 2<sup>e</sup> année. La distribution pour les filles de 2<sup>e</sup> année est aussi significativement leptokurtique. Notons que ce groupe de filles se distingue également en ce qui concerne leur niveau cognitif atteint. On constate également une diminution progressive du réalisme moyen des élèves dans les niveaux scolaires plus élevés. La nature de ce phénomène sera examinée ultérieurement.

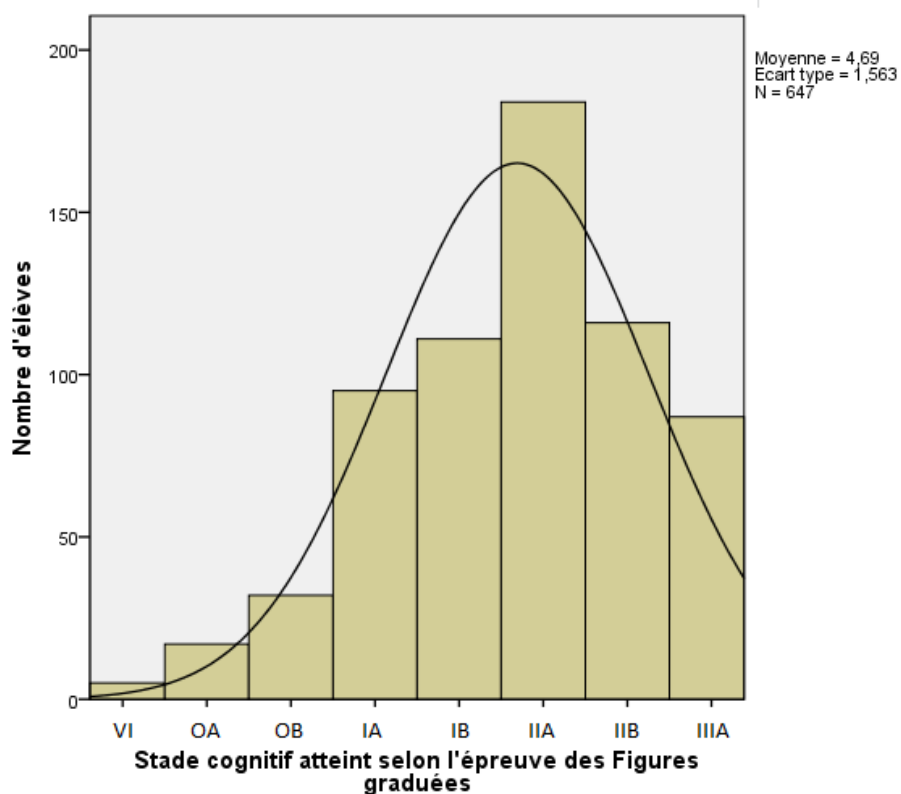


Figure 9. Distribution des stades atteints par les élèves selon l'épreuve des Figures graduées.

Tableau 11

*Statistiques sur les distributions des stades cognitifs atteints selon l'épreuve des Figures graduées*

	N	Moy.	É-T	Asymétrie	Erreur type	z	Kurtose	Erreur type	z
Filles du niveau scolaire 2	124	3,90	1,464	-,574	,217	-2,640 *	-,019	,431	-0,045
Filles du niveau scolaire 5	105	4,74	1,264	-,579	,236	-2,455	,348	,467	0,745
Filles du niveau scolaire 8	119	5,70	1,273	-,817	,222	-3,682 *	,117	,440	0,266
Garçons du niveau scolaire 2	109	3,39	1,305	-,210	,231	-0,907	,099	,459	0,216
Garçons du niveau scolaire 5	91	4,87	1,301	-,647	,253	-2,560	,525	,500	1,048
Garçons du niveau scolaire 8	99	5,69	1,184	-,722	,243	-2,977 *	,052	,481	0,109

Légende : Les stades ont été recodés sur une échelle numérique pour des fins de calculs. Par exemple, VI = 0, OA = 1, etc.

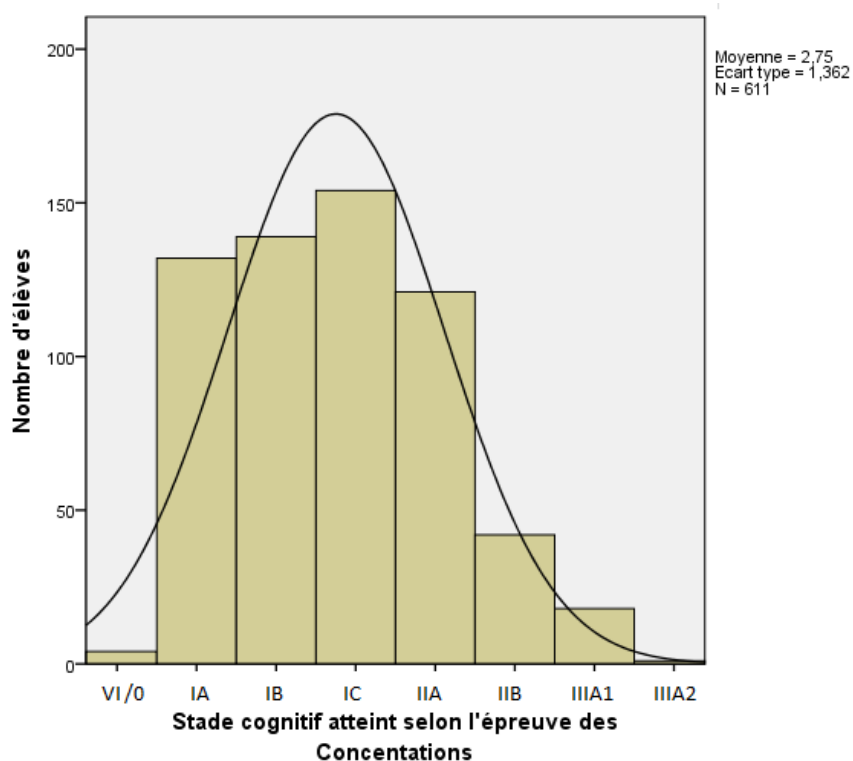


Figure 10. Distribution des stades atteints pour l'ensemble des élèves selon l'épreuve des Concentrations.

Tableau 12

*Statistiques sur les distributions des stades cognitifs atteints selon l'épreuve des Concentrations*

	N	Moy.	É-T	Asymétrie	Erreur type	z	Kurtose	Erreur type	z
Filles du niveau scolaire 2	111	1,78	,814	,315	,229	1,375	-1,217	,455	-2,673 *
Filles du niveau scolaire 5	103	2,58	1,005	,094	,238	0,395	,354	,472	0,750
Filles du niveau scolaire 8	115	3,72	1,189	,048	,226	0,211	-,041	,447	-0,092
Garçons du niveau scolaire 2	98	1,77	,859	,575	,244	2,361	-,648	,483	-1,341
Garçons du niveau scolaire 5	89	2,52	1,129	-,067	,255	-0,262	-1,019	,506	-2,015
Garçons du niveau scolaire 8	95	4,13	1,113	-,444	,247	-1,793	,382	,490	0,779

Légende : Les stades ont été recodés sur une échelle numérique pour des fins de calculs. Par exemple, VI/O = 0, IA = 1, etc.

La Figure 12 montre la distribution du réalisme des élèves pour l'épreuve des Concentrations, et ce, pour l'ensemble des élèves, soit les élèves de la 2<sup>e</sup>, 5<sup>e</sup>, et 8<sup>e</sup> année. Nous constatons une légère asymétrie positive. Le Tableau 14 fait état des coefficients d'asymétrie et de kurtose pour les élèves de chaque niveau scolaire tout en séparant les garçons et les filles. Nous constatons que les coefficients d'asymétrie sont presque tous positifs et significatifs dans le cas particulier des filles et des garçons de la 2<sup>e</sup> année.

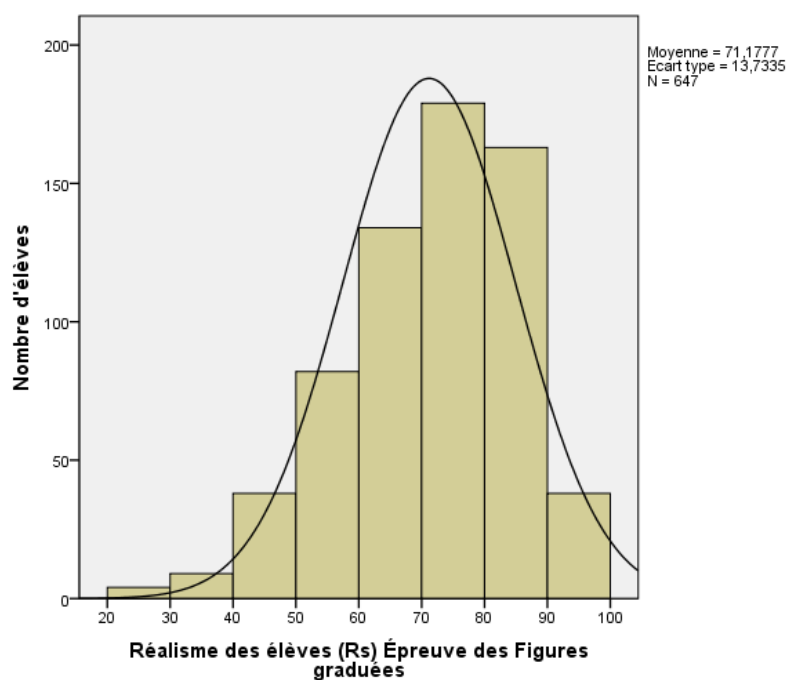


Figure 11. Distribution du réalisme pour l'ensemble des élèves avec l'épreuve des Figures graduées.

Tableau 13

Statistiques sur les distributions du réalisme des élèves avec l'épreuve des Figures graduées

	N	Moy.	É-T	Asymétrie	Erreur type	z	Kurtose	Erreur type	z
Filles du niveau scolaire 2	124	75,923	11,627	-1,034	,217	-4,755 *	1,398	,431	3,239 *
Filles du niveau scolaire 5	105	73,730	11,869	-,353	,236	-1,496	-,394	,467	-,844
Filles du niveau scolaire 8	119	67,101	14,617	-,438	,222	-1,977	-,263	,440	-,597
Garçons du niveau scolaire 2	109	68,149	15,055	-,260	,231	-1,124	-,765	,459	-1,666
Garçons du niveau scolaire 5	91	75,991	10,606	-,609	,253	-2,410	-,144	,500	-,287
Garçons du niveau scolaire 8	99	66,338	14,238	-,374	,243	-1,541	-,126	,481	-,262

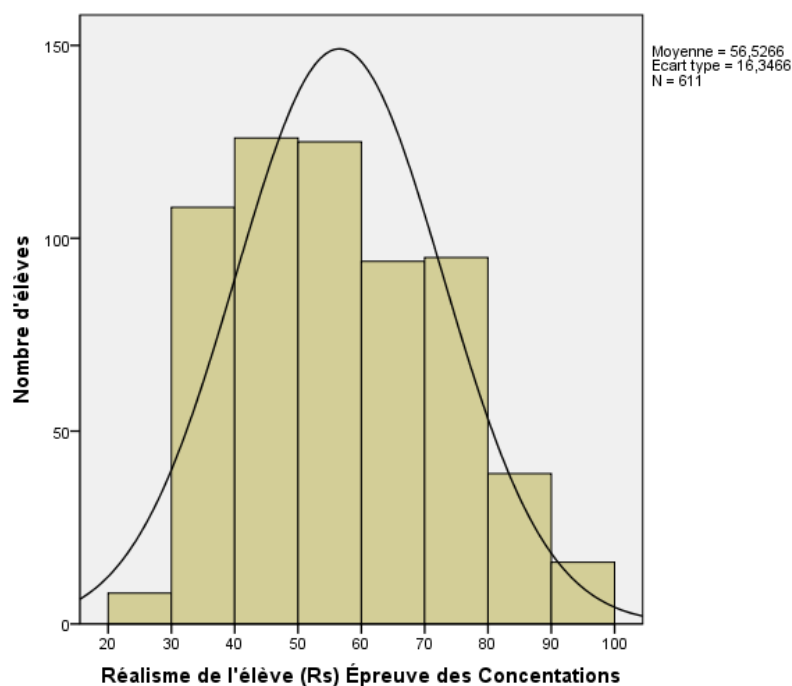


Figure 12. Distribution du réalisme pour l'ensemble des élèves avec l'épreuve des Concentrations.

Tableau 14

Statistiques sur les distributions du réalisme des élèves avec l'épreuve des Concentrations

	N	Moy.	É-T	Asymétrie	Erreur type	z	Kurtose	Erreur type	z
Filles du niveau scolaire 2	111	48,104	12,517	0,806	0,229	3,512 *	-0,286	0,455	-,629
Filles du niveau scolaire 5	103	56,255	14,347	0,058	0,238	,245	-0,613	0,472	-1,299
Filles du niveau scolaire 8	115	69,274	13,554	-0,080	0,226	-,356	-0,380	0,447	-,850
Garçons du niveau scolaire 2	98	44,155	11,448	0,929	0,244	3,812 *	0,440	0,483	,912
Garçons du niveau scolaire 5	89	50,512	12,007	0,544	0,255	2,130	-0,414	0,506	-,819
Garçons du niveau scolaire 8	95	69,627	13,580	-0,188	0,247	-,760	-0,649	0,490	-1,324

## VÉRIFICATION DES QUESTIONS DE RECHERCHE

### Postulats

Des analyses de type multivarié ont été effectuées. Ce type d'analyse postule que les distributions doivent être conformes à certains postulats concernant la normalité, l'homoscédasticité et la linéarité. Les postulats de normalité, de linéarité et d'homoscédasticité ont été vérifiés pour les variables suivantes : réalisme des élèves pour l'épreuve des Figures graduées et pour l'épreuve des Concentrations. Le test de Kolmogorov-Smirnov n'a pas révélé d'anormalités significatives dans les distributions.

Les Figures 13 et 14 montrent, dans le cas de l'épreuve des Figures graduées, qu'il y a absence de linéarité et d'homoscédasticité chez les garçons comme chez les filles. Nous verrons un peu plus loin qu'il s'agit d'un phénomène de plafonnement. Nous devons en tenir compte dans nos analyses. Par ailleurs, en ce qui concerne l'épreuve des Concentrations, la conformité aux postulats de linéarité et d'homoscédasticité est assurée. Les Figures 15 et 16 en font état avec des  $R^2$  linéaires de 0,528 et 0,578 respectivement.

Enfin, rappelons que le calcul du réalisme utilise les degrés de certitude et la qualité des réponses. La détermination des stades atteints utilise également la qualité des réponses. Ceci peut élever le risque de multicollinéarité entre certaines variables, notamment des calculs de régression multiple. La multicollinéarité a donc dû être vérifiée.

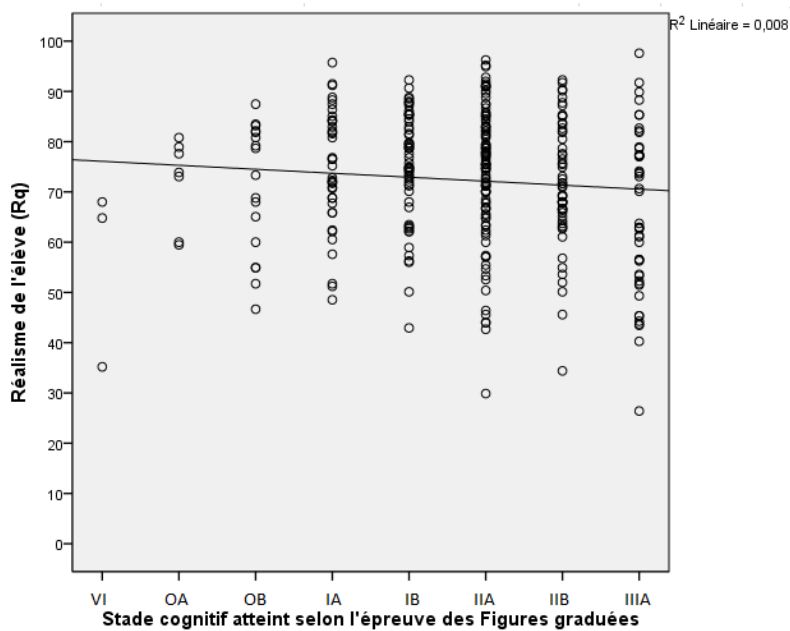


Figure 13. Relation entre le réalisme des filles et le stade cognitif atteint selon l'épreuve des Figures graduées.

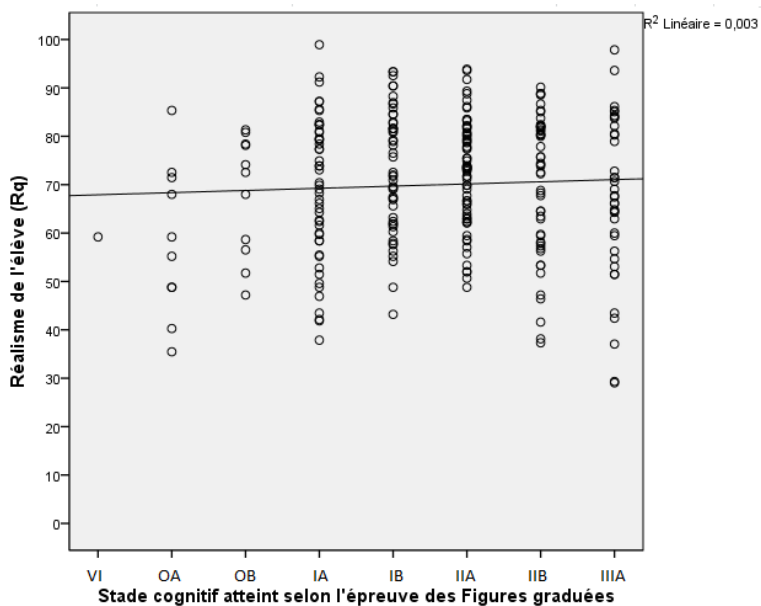


Figure 14. Relation entre le réalisme des garçons et le stade cognitif atteint selon l'épreuve des Figures graduées.

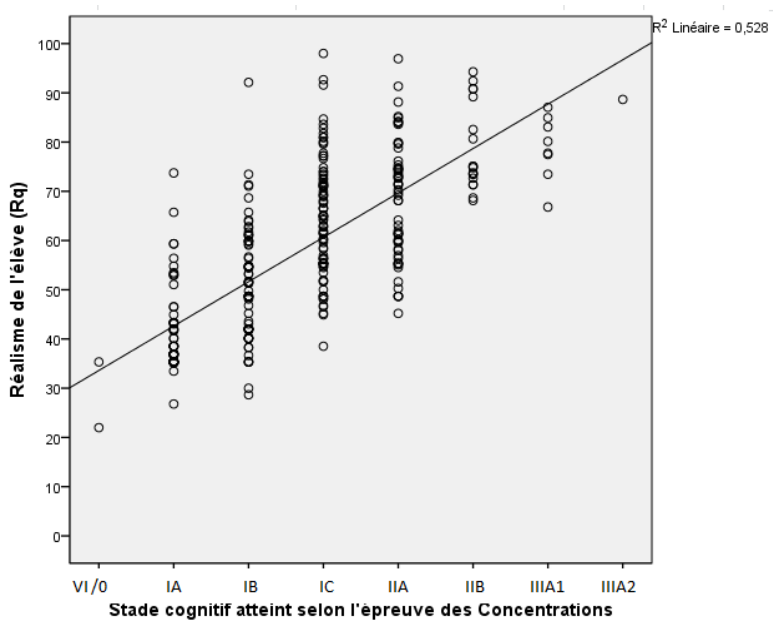


Figure 15. Relation entre le réalisme des filles et le stade cognitif atteint selon l'épreuve des Concentrations.

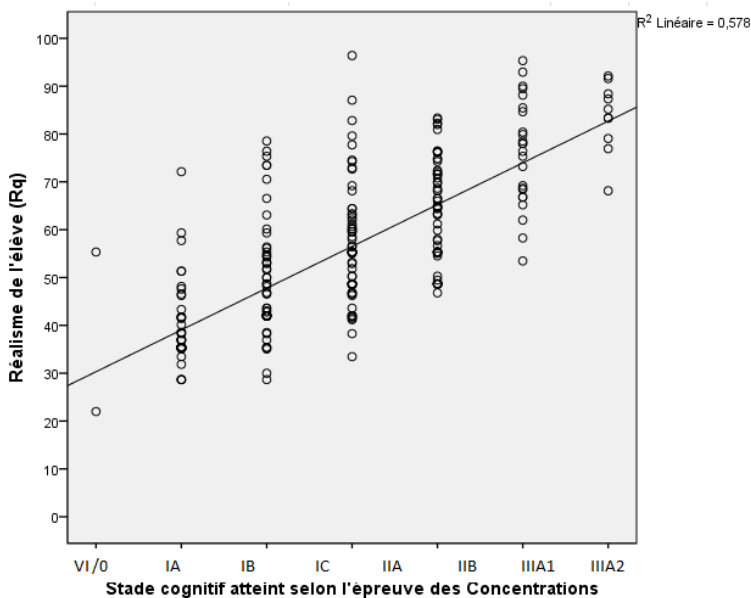


Figure 16. Relation entre le réalisme des garçons et le stade cognitif atteint selon l'épreuve des Concentrations.

### Liens entre les variables – régression multiple

La première étape des analyses de régression multiple a été de vérifier s'il y a multicolinéarité entre les variables indépendantes par rapport à la variable dépendante. L'absence de colinéarité entre les variables est une des conditions d'utilisation de la régression multiple. La colinéarité entre les variables se produit lorsque deux ou plusieurs variables indépendantes sont redondantes les unes par rapport aux autres. Des variables peuvent varier proportionnellement dans le même sens et causer une surestimation de certains effets. Par exemple, les notes en lecture et en écriture d'un élève au secondaire constituent d'excellents prédicteurs des résultats en français au collège ou à l'université. Les notes en lecture et en écriture d'un élève sont fortement corrélées entre elles et sont colinéaires dans le cas présenté ici. On doit utiliser qu'une seule des deux pour éviter de surestimer le potentiel prédictif des notes par rapport au niveau de performance ultérieure de l'élève. Les Tableaux 15 et 16 montrent les coefficients VIF et de Tolérance pour les variables des deux épreuves. Aucun coefficient VIF n'excède le seuil de 10 et aucun coefficient de tolérance ne s'approche de 0. On peut donc conclure que les variables sont conformes à cette condition d'utilisation.

Les Tableaux 15 et 16 montrent les résultats des analyses de régression multiple linéaire pour les deux épreuves. Le but principal de ces analyses est de départager la proportion de la variance du réalisme attribuable aux différentes variables indépendantes, soit le stade cognitif atteint selon l'épreuve, le niveau scolaire et le sexe de l'élève. Le  $R$  carré global nous indique la proportion de la variance expliquée de la variable dépendante par les variables indépendantes combinées. Même si le modèle de régression linéaire est significatif pour l'épreuve des Figures graduées,  $R$  carré global ( $R^2 = 0,045$ ) indique que la variance expliquée est très faible. Par contre,

pour l'épreuve des Concentrations, 57,6 % de la variance du réalisme ( $R^2 = 0,576$ ) est expliquée par le stade cognitif atteint, le niveau scolaire et le sexe de l'élève. Voici l'équation dans le cas de l'épreuve des Concentrations. L'équation permet de constater que pour chaque changement de 1 point du stade atteint, le réalisme de l'élève change de 7,426 points lorsque les autres variables indépendantes demeurent constantes.

$$\text{Réalisme}(Rs) = (35,584 + 7,426 \times \text{Stade} + 1,121 \times \text{Niveau scolaire} - 3,834 \times \text{Sexe})$$

Dans le cas de l'épreuve des Figures graduées, nous remarquons qu'aucune variable indépendante ne se démarque en matière de corrélation avec la variable dépendante qu'est le réalisme de l'élève. Les coefficients de corrélation partielle sont beaucoup plus élevés pour l'épreuve des Concentrations avec ( $r = 0,586$ ) pour le stade cognitif atteint ( $r = 0,212$ ) pour le niveau scolaire et ( $r = -0,178$ ) pour le sexe de l'élève. Le stade cognitif atteint et le niveau scolaire sont plus corrélés avec le réalisme de l'élève que le sexe. Les analyses de régression subséquentes vont permettre de constater les effets de ces corrélations partielles.

Tableau 15

*Résumé du modèle proposé par l'analyse de la régression pour le réalisme (Rs) de l'épreuve des Figures graduées*

Résumé du modèle de la régression multiple

Modèle	R	R carré	F	Sig.
Score du sujet	0,213	0,045	10,193	0,000

	Coefficients bruts		Standardisés			Corrélations		Statistiques de colinéarité	
	B	Erreur standard	Beta	t	Sig.	Corrélation simple	Partielle	Tolérance	VIF
Ordonnée à l'origine	75,585	2,348		32,193	,000				
Stade atteint	1,091	,404	,124	2,701	,007	,002	,106	,702	1,424
Niveau scolaire	-1,279	,252	-,233	-5,083	,000	-,165	-,197	,704	1,420
Sexe	-2,201	1,062	-,080	-2,072	,039	-,084	-,081	,997	1,003

Tableau 16

*Résumé du modèle proposé par l'analyse de la régression pour le réalisme (Rs) de l'épreuve des Concentrations*

Résumé du modèle de la régression multiple

Modèle	R	R carré	F	Sig.
Score du sujet	0,759	0,576	274,707	0,000

	Coefficients bruts		Standardisés			Corrélations		Statistiques de colinéarité	
	B	Erreur standard	Beta	t	Sig.	Corrélation simple	Partielle	Tolérance	VIF
Ordonnée à l'origine	35,584	1,631		21,822	,000				
Stade atteint	7,426	,417	,619	17,810	,000	,735	,586	,579	1,728
Niveau scolaire	1,221	,228	,186	5,346	,000	,588	,212	,579	1,726
Sexe	-3,834	,862	-,118	-4,449	,000	-,099	-,178	,997	1,003

Liens entre les variables – corrélation entre les variables dépendantes

Le Tableau 17 ci-après montre les corrélations non paramétriques entre le réalisme des élèves et les stades atteints dans les deux épreuves. Dans le cas de l'épreuve des Concentrations, les stades atteints et le réalisme des élèves sont corrélés significativement ( $r = 0,747$ ;  $p < 0,000$ ). Il n'en est pas de même pour l'épreuve des Figures graduées ( $r = -0,023$ ;  $p = 0,551$ ). Cette

absence de lien peut être constatée avec les Figures 19 et 20 présentées ultérieurement. Il semble y avoir un effet de plafonnement autour du niveau de réalisme correspondant à 0,71.

Nous constatons également des liens significatifs entre les stades atteints selon les deux épreuves ( $r = 0,447$ ;  $p < 0,000$ ) ainsi qu'entre les stades atteints à l'épreuve des Figures graduées et le réalisme des élèves à l'épreuve des Concentrations ( $r = 0,395$ ,  $p < 0,000$ ). Visiblement, les deux épreuves permettent de déterminer le stade atteint. Il est donc normal que le stade atteint déterminé par l'épreuve des Figures graduées soit corrélé avec le réalisme déterminé par l'épreuve des Concentrations puisque le réalisme est corrélé avec cette épreuve.

Nous avons également vérifié la corrélation entre le réalisme des élèves d'une épreuve à l'autre ( $r = -,167$ ,  $p < 0,000$ ). Les Figures 17 et 18 montrent le lien de linéarité entre les deux épreuves. Nous constatons un faible lien linéaire ou curvilinéaire avec des coefficients de détermination qui sont très faibles. Ces valeurs sont négligeables et ne feront pas l'objet d'une interprétation particulière. Notons néanmoins que l'effet de plafonnement observé dans le cas de l'épreuve des Figures graduées y est pour quelque chose.

Tableau 17

*Corrélation entre les stades déterminés par les deux épreuves et le réalisme des élèves*

Rho de Spearman		FG-STADE	Conc-STADE	FG-Rs	Conc-Rs
FG-STADE	Coefficient de corrélation	1,000			
	Sig. (bilatérale)	.			
	N	650			
Conc-STADE	Coefficient de corrélation	,447**	1,000		
	Sig. (bilatérale)	,000	.		
	N	628	636		
FG-Rs	Coefficient de corrélation	-,023	-,123**	1,000	
	Sig. (bilatérale)	,551	,002	.	
	N	647	625	647	
Conc-Rs	Coefficient de corrélation	,395**	,747**	-,167**	1,000
	Sig. (bilatérale)	,000	,000	,000	.
	N	603	611	601	611

Légende :

FG-Stade correspond au stade cognitif atteint tel que déterminé par l'épreuve des Figures graduées

Conc-Stade correspond au stade cognitif atteint tel que déterminé par l'épreuve des Concentrations

FG-Rs correspond au réalisme de l'élève pour l'épreuve des Figures graduées

Conc-Rs correspond au réalisme de l'élève pour l'épreuve des Concentrations

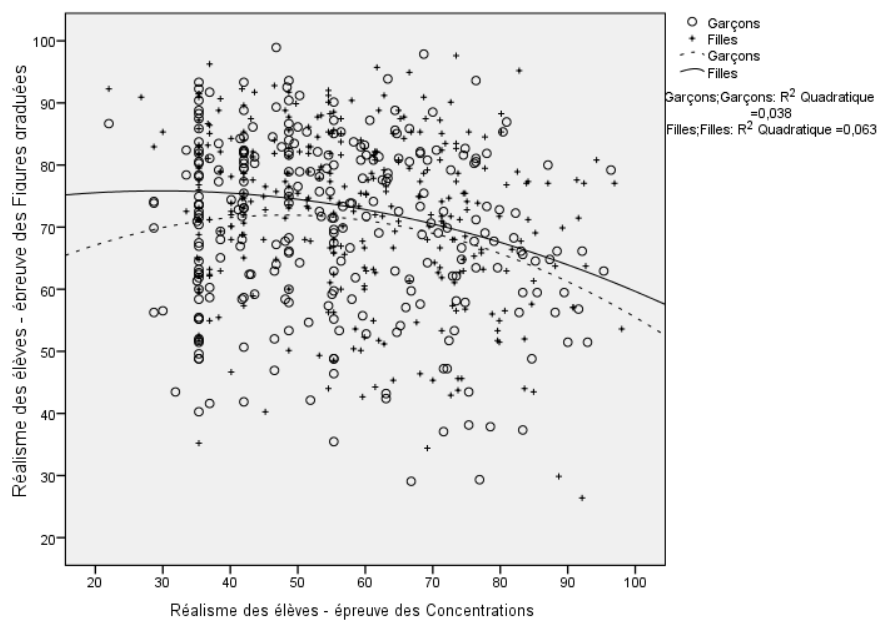


Figure 17. Relation curvilinéaire entre le réalisme des élèves en fonction des deux épreuves.

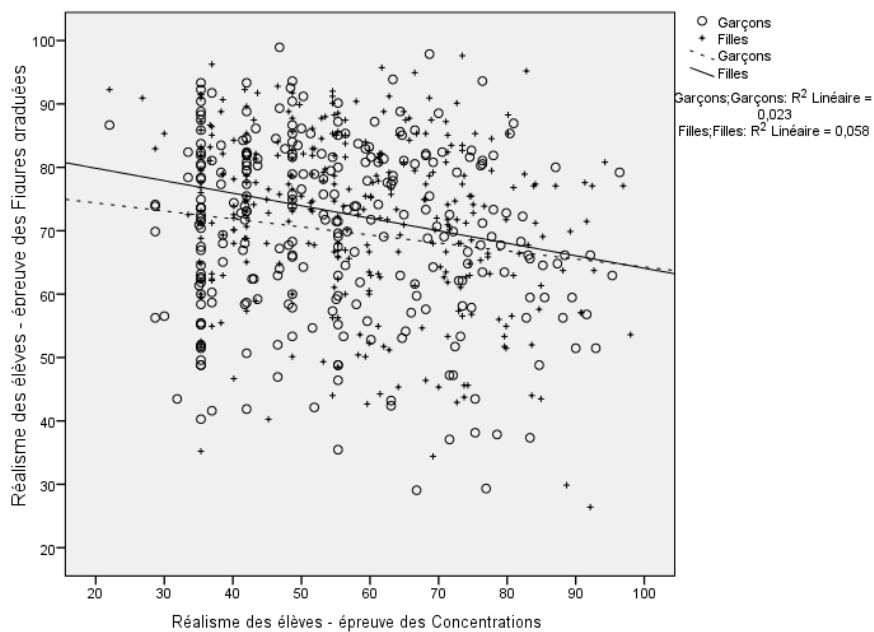


Figure 18. Relation linéaire entre le réalisme des élèves en fonction des deux épreuves.

### **Effet du stade cognitif atteint sur le réalisme des élèves**

Une des questions de recherche vise à déterminer l'âge minimal requis pour que le réalisme des élèves puisse être utilisé de façon valide dans un contexte scolaire pour la majorité des élèves. Nous avons préféré employer le stade cognitif atteint par l'élève plutôt que son âge puisque celui-ci n'est pas garant de l'atteinte d'un stade cognitif en particulier, malgré un lien marqué entre l'âge et les stades cognitifs. En d'autres mots, à compter de quel stade cognitif atteint par les élèves, l'utilisation de l'indice de réalisme des élèves devient-elle pertinente?

Pour effectuer une telle étude, la moyenne du réalisme des élèves est analysée en fonction du stade cognitif atteint. La variable sexe est contrôlée de manière à ne pas confondre certains effets. Pour l'épreuve des Figures graduées et l'épreuve des Concentrations, nous avons effectué des analyses séparées; les Figures 19 à 22 montrent le réalisme moyen des élèves à chaque stade, dans chaque épreuve et pour chaque sexe.

Le niveau de réalisme semble progresser en fonction du stade cognitif atteint et plafonne autour du niveau de réalisme correspondant à 0,71. Ceci est très visible dans le cas de l'épreuve des Figures graduées, et ce, indépendamment du sexe de l'élève. En ce qui concerne l'épreuve des Concentrations, le niveau de réalisme semble progresser en fonction du stade cognitif atteint de façon quasi linéaire sans égard au sexe. La matrice de corrélation au Tableau 17 en fait d'ailleurs état.



Figure 19. Réalisme moyen des filles en fonction du stade atteint selon l'épreuve des Figures graduées.

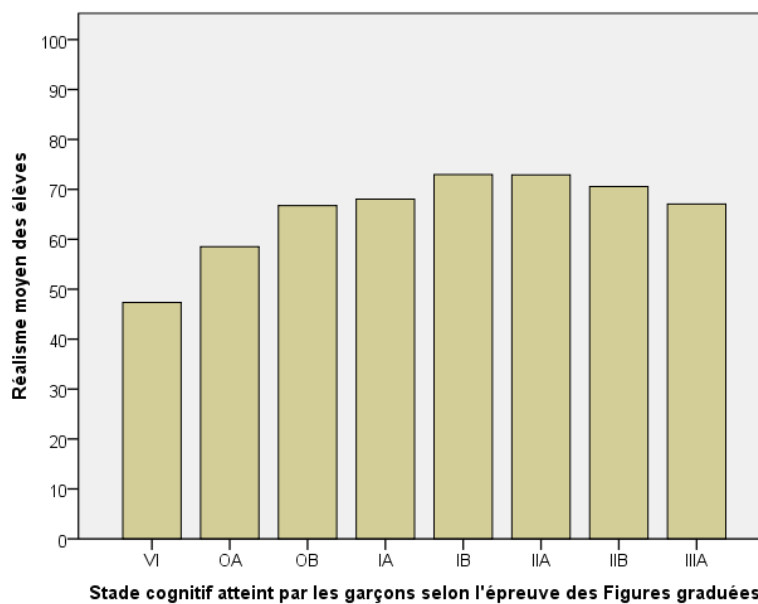


Figure 20. Réalisme moyen des garçons en fonction du stade atteint selon l'épreuve des Figures graduées.

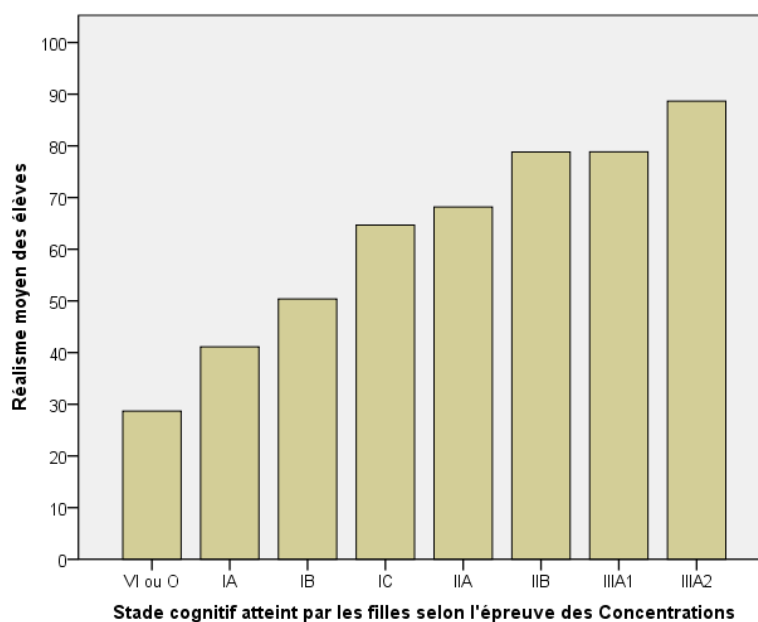


Figure 21. Réalisme moyen des filles en fonction du stade atteint selon l'épreuve des Concentrations.

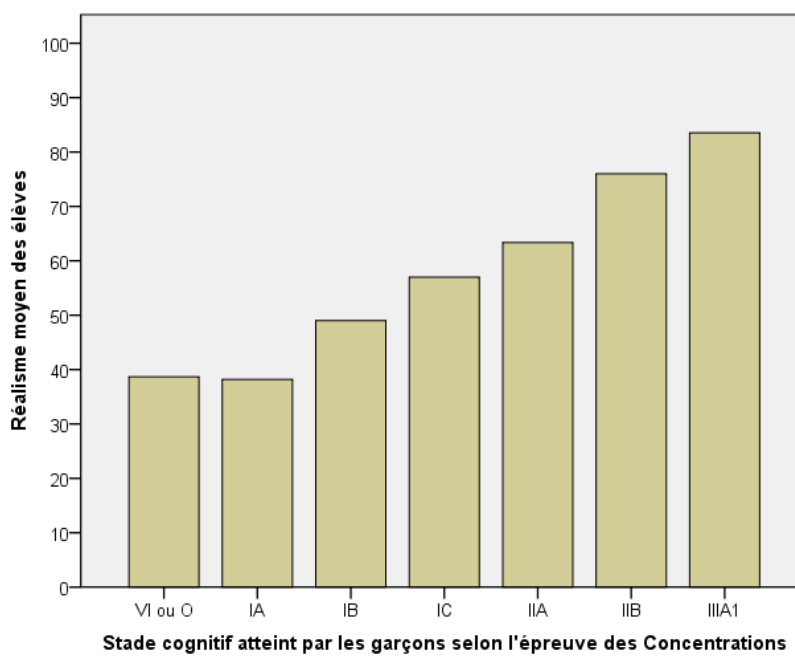


Figure 22. Réalisme moyen des garçons en fonction du stade atteint selon l'épreuve des Concentrations.

De manière à déterminer le stade cognitif minimal requis pour rendre pertinente l'utilisation de l'indice de réalisme, nous avons dû nous appuyer sur une norme permettant d'identifier un point de césure qui discrimine entre le réalisme faible et le réalisme insuffisant. À cause du manque de normes locales, nous nous sommes référés aux recommandations de Leclercq (1993) qui sont validées par la suite par Gilles (1996). Le Tableau 3 dans la section qui porte sur la recension des écrits présente les normes établies à partir de vingt-huit examens administrés par la Faculté de psychologie et des sciences sociales de l'Université de Liège. Le point de césure à 70 % correspond approximativement au point d'inflexion inférieur de la courbe normale, c'est-à-dire au point limite qui inclut 30 % des cas les plus faibles en matière de réalisme. Dans le cas de l'échantillon de Gilles, 25,4 % des étudiants se situaient sous la barre du 70 % d'indice de réalisme. Notons que le point d'inflexion inférieur de la courbe normale se situe à 27 %. Si nous acceptons provisoirement que le point de césure de la pertinence de l'indice de réalisme se situe autour de 70 %, nous obtenons le stade cognitif minimal pour qu'il soit pertinent de considérer le jugement des élèves dans l'auto-estimation de la qualité de leurs réponses en situation de *testing*. En ce qui concerne l'épreuve des Figures graduées, nous remarquons que l'atteinte du niveau de réalisme correspondant à environ 0,71 se produit chez les filles ayant atteint les niveaux de développement supérieurs à OA avec un réalisme (*Rs*) moyen de 0,72 et OB chez les garçons avec un réalisme (*Rs*) moyen de 0,68. Pour l'épreuve des Concentrations, le seuil se situe au niveau de développement IIA pour les filles avec un réalisme (*Rs*) moyen de 0,68 et quelque part entre IIA et IIB chez les garçons. Il apparaît que les élèves en deuxième année commencent à être suffisamment réalistes en ce qui concerne les tâches de l'épreuve des Figures graduées. Nous notons que, chez les garçons, le stade requis est un peu

plus élevé que chez les filles dans les deux épreuves. Du côté de l'épreuve des Concentrations, il semble que le réalisme prenne plus de temps à se développer pour les filles et les garçons. Le Tableau 18 rend compte des stades cognitifs minimaux requis selon le sexe et l'épreuve. Par exemple, nous constatons par les figures 19 et 20 que les filles atteignent le seuil minimal de réalisme au stade OA alors que les garçons l'atteignent au stade OB dans le cas de l'épreuve des Figures graduées. Notons également que même si le stade cognitif minimal requis pour rendre pertinente l'utilisation de l'indice de réalisme n'est pas atteint, nous pouvons néanmoins l'introduire en postulant que la métacognition doit se développer chez l'élève.

Tableau 18

*Stades cognitifs requis pour que le réalisme représente une valeur ajoutée*

	Spatio-temporel	Logico-mathématique
Filles	OA-Symbolique Inférieur	IIA-Opératoire Concret Inférieur
Garçons	OB-Symbolique Supérieur	IIA-Opératoire Concret Inférieur

### **Réalisme des élèves en fonction du niveau scolaire et du sexe**

Nous avons examiné l'effet du niveau scolaire et du sexe des élèves sur le réalisme dans les deux épreuves. Une analyse multivariée de la variance calculée avec SPSSv.19 (Modèle linéaire général) a permis d'en apprendre plus sur ces liens et sur les interactions

potentielles entre les variables indépendantes. Les Tableaux 19 et 20 qui suivent présentent certaines statistiques descriptives sur le réalisme de l'élève utilisé à titre de variable dépendante pour l'analyse multivariée de la variance de chaque épreuve. Les statistiques présentées sont en l'occurrence la moyenne, l'écart-type et la variance du réalisme en fonction des variables indépendantes telles le sexe et le niveau scolaire de l'élève. Nous notons que la variance du réalisme des élèves semble se distribuer inégalement.

Tableau 19

*Statistiques descriptives sur l'analyse multivariée de la variance de réalisme pour l'épreuve des Figures graduées*

Sexe	Niveau	Moyenne	Écart-type	Variance
F	2	75,86	11,37	135,19
	5	73,90	11,89	140,88
	8	67,59	14,51	213,66
G	2	67,97	14,70	226,66
	5	76,01	10,78	112,49
	8	66,57	14,15	202,73
Total	2	72,18	13,58	192,30
	5	74,85	11,42	128,33
	8	67,13	14,33	207,88

Tableau 20

*Statistiques descriptives sur l'analyse multivariée de la variance de réalisme pour l'épreuve des Concentrations*

Sexe	Niveau	Moyenne	Écart-type	Variance
F	2	48,10	12,57	156,68
	5	56,43	14,31	205,84
	8	69,27	13,55	183,72
G	2	44,18	11,53	131,06
	5	50,57	12,27	144,17
	8	69,68	13,61	184,41
Total	2	46,27	12,23	147,88
	5	53,76	13,70	184,59
	8	69,46	13,55	183,19

Le Tableau 21 présente les résultats de l'analyse multivariée de la variance du réalisme des élèves pour les deux épreuves. Nous constatons que les deux effets principaux sont significatifs. En utilisant le Lambda de Wilks à titre de critère, nous constatons une différence significative du réalisme moyen entre les garçons et les filles ( $F = 7,267$  ;  $p = 0,001$ ) et leur niveau scolaire ( $F = 81,970$  ;  $p < 0,000$ ). Il faut aussi noter une interaction significative entre le sexe et le niveau scolaire ( $F = 5,320$  ;  $p < 0,000$ ). Notons que le devis ainsi que le nombre d'élèves ont permis d'avoir suffisamment de puissance avec respectivement 93,6 %, 100 % et 97,3 % de probabilité de rejeter les hypothèses nulles lorsqu'elles étaient fausses. Les valeurs d'êta-carré sont assez faibles pour le sexe (0,024) et l'interaction (0,018), ce qui indique qu'une faible proportion de la variance est expliquée par les différents niveaux de ces conditions. La valeur d'êta-carré pour le niveau scolaire (0,216) est plus importante. Puisque les liens sont

significatifs, des analyses univariées sont recommandées et sont présentées dans la prochaine section.

Nous avons déjà observé avec les Figures 13 et 14 que la distribution des résultats n'est pas nécessairement conforme aux postulats de linéarité et d'homoscédasticité. Le manque de conformité au postulat d'homoscédasticité montre que les variables indépendantes ont peu ou pas de liens clairs avec le réalisme des élèves. Le test de Box confirme cette observation ( $F = 2,119$  ;  $p = 0,007$ ). L'analyse Roy-Bargmann ne semble pas indiquée puisque sa performance peut être affectée entre autres par le manque de conformité au postulat d'homoscédasticité.

Tableau 21

*Tableau résumé de l'analyse multivariée de la variance de réalisme pour les deux épreuves*

Source		Valeur	<i>DI</i>	<i>F</i>	Sig.	Êta-carré partiel	Puissance
Sexe	Pillai	,024	2,000	7,267 <sup>a</sup>	,001	,024	,936
	Wilks	,976	2,000	7,267 <sup>a</sup>	,001	,024	,936
	Hotelling	,024	2,000	7,267 <sup>a</sup>	,001	,024	,936
	Roy	,024	2,000	7,267 <sup>a</sup>	,001	,024	,936
Niveau scolaire	Pillai	,395	4,000	73,182	,000	,197	1,000
	Wilks	,614	4,000	81,970 <sup>a</sup>	,000	,216	1,000
	Hotelling	,613	4,000	90,942	,000	,235	1,000
	Roy	,588	2,000	175,055 <sup>c</sup>	,000	,370	1,000
Sexe*niveau scolaire	Pillai	,035	4,000	5,320	,000	,018	,973
	Wilks	,965	4,000	5,320 <sup>a</sup>	,000	,018	,973
	Hotelling	,036	4,000	5,320	,000	,018	,973
	Roy	,026	2,000	7,677 <sup>c</sup>	,001	,025	,948

Note : Les variables indépendantes sont des variables catégorielles traitées à titre de facteurs fixes. Les variables dépendantes sont l'indice de réalisme pour chaque épreuve.

Tableau 22

*Tableau résumé des analyses univariées de la variance du réalisme pour FG et Conc*

Source		SC	Dl	CM	F	Sig.	Êta-carré partiel	Puissance
Sexe	FG	764,596	1	764,596	4,501	,034	,008	,563
	CONC	1449,544	1	1449,544	8,523	,004	,014	,830
Niveau scolaire	FG	6207,638	2	3103,819	18,272	,000	,058	1,000
	CONC	58321,458	2	29160,729	171,455	,000	,366	1,000
Sexe*niveau scolaire	FG	2586,056	2	1293,028	7,612	,001	,025	,946
	CONC	1020,409	2	510,205	3,000	,051	,010	,582
Erreur	FG	101069,645	595	169,865				
	CONC	101196,485	595	170,078				
Total	FG	3162974,776	601					
	CONC	2088715,569	601					

Alpha = 0,05      R carré = 0,087 pour FG et R carré = 0,368 pour FG

Note : Les variables indépendantes sont des variables catégorielles traitées à titre de facteurs fixes. Les variables dépendantes sont l'indice de réalisme pour chaque épreuve, soit FG pour Figures graduées et CONC pour Concentrations.

### Épreuve des Figures graduées

Le Tableau 22 présente les résultats de l'analyse univariée de la variance du réalisme des élèves pour les deux épreuves. Nous constatons une différence significative du réalisme moyen entre les garçons et les filles ( $F = 4,501$  ;  $p = 0,034$ ) et leur niveau scolaire ( $F = 18,272$  ;  $p < 0,000$ ). Dans le cas précis de l'épreuve des Figures graduées, il faut aussi noter une interaction significative entre le sexe et le niveau scolaire ( $F = 7,612$  ;  $p = 0,001$ ). Dans le cas de l'épreuve des Concentrations, nous n'observons pas d'interaction significative entre le sexe et le niveau scolaire. Notons que le devis ainsi que le nombre d'élèves ont permis d'avoir suffisamment de puissance avec respectivement 56,3 %, 100 % et 94,6 % de probabilité de rejeter les hypothèses nulles lorsqu'elles étaient fausses. Les valeurs d'êta-carré sont assez faibles pour le sexe (0,008), le niveau scolaire (0,058) et l'interaction (0,025), ce qui indique qu'une faible proportion de la variance est expliquée par les différents niveaux de ces conditions. Le Tableau 22 montre que l'interaction est significative tant entre les élèves de la 2<sup>e</sup> année à la 5<sup>e</sup> année ( $p = 0,022$ ) que les élèves de la 5<sup>e</sup> année à la 8<sup>e</sup> année ( $p < 0,000$ ).

Le  $R$  carré global avec 8,7 % de la variance du réalisme expliquée par les variables indépendantes combinées est plutôt faible et suggère que les liens entre les variables indépendantes et la variable dépendante sont assez faibles. Nous avons déjà observé avec les Figures 13 et 14 que la distribution des résultats n'est pas nécessairement conforme aux postulats de linéarité et d'homoscédasticité. Le manque de conformité au postulat d'homoscédasticité montre que les variables indépendantes ont peu ou pas de lien clair avec le réalisme des élèves.

Le test d'homogénéité de la variance de Levene révèle que la variance est inégale entre les différents groupes ( $F = 3,950$  ;  $p = 0,002$ ). Lorsqu'on vérifie l'homogénéité de la variance pour chaque variable, le niveau scolaire manque de conformité au postulat d'homoscédasticité de façon significative ( $p = 0,020$ ) alors que ce n'est pas significatif pour le sexe ( $p = 0,900$ ). La variance du réalisme se distribue de la façon suivante ( $2^e = 192,30$ ;  $5^e = 128,33$ ;  $8^e = 207,88$ ). La distribution des résultats n'est pas conforme aux postulats de linéarité et d'homoscédasticité. La violation du postulat d'homoscédasticité fait en sorte que les analyses ne permettent pas de démontrer l'existence d'un lien entre les variables indépendantes et le réalisme des élèves.

Tableau 23

*Résultats des contrastes pour les deux épreuves*

		FG	CONC
Niveau 1 et niveau 2	Estimation de contraste	-3,039	-7,357
	Différence (estimation - hypothèse)	-3,039	-7,357
	Erreur	1,321	1,322
	Sig.	,022	,000
Niveau 2 et niveau 3	Estimation de contraste	7,873	-15,981
	Différence (estimation - hypothèse)	7,873	-15,981
	Erreur	1,320	1,321
	Sig.	,000	,000

Légende : Niveau 1, 2 et 3 réfèrent aux niveaux scolaires correspondant à la 2<sup>e</sup> année, 5<sup>e</sup> année et 8<sup>e</sup> année respectivement. Il s'agit de contrastes simples qui visent à montrer les différences entre les niveaux. FG et CONC réfèrent à Figures graduées et Concentrations respectivement.

L'analyse univariée de la variance a permis de constater qu'il y a une différence significative entre les moyennes de réalisme des filles et des garçons en général. À la lumière de ces résultats, nous pouvons rejeter l'hypothèse nulle concernant l'effet principal lié au sexe de l'élève. Il apparaît donc que le réalisme des élèves est lié directement au sexe de l'élève dans les domaines spatio-temporel. Nous rejetons également l'hypothèse nulle concernant l'effet principal lié au niveau scolaire ou à l'âge de l'élève. Les résultats soutiennent que le réalisme de l'élève varie en fonction de son niveau scolaire. L'interaction entre les deux conditions étudiées est significative. Nous notons néanmoins que les analyses sont affaiblies par le manque de conformité de certains postulats. Les effets expérimentaux sont également très faibles ainsi que les valeurs de puissance. Ceci est confirmé par les résultats de la régression multiple.

La Figure 23 permet de mieux visualiser l'interaction entre le réalisme des élèves et leur niveau scolaire pour l'épreuve des Figures graduées. Nous constatons une différence entre le réalisme moyen des filles et celui des garçons dans les tâches de l'épreuve des Figures graduées. Les résultats indiquent aussi que le réalisme des filles et des garçons varie entre les niveaux scolaires 5 et 8. En fait, nous constatons une tendance baissière du réalisme de l'élève avec le niveau scolaire. La différence de réalisme entre les garçons et les filles de 2<sup>e</sup> année explique majoritairement l'interaction observée.

Rappelons que selon les normes de Gilles (1996), le seuil minimal pour qu'un étudiant adulte soit considéré minimalement réaliste se situe autour de 0,71. Dans le cas de l'épreuve de Figures graduées, à peu près tous les élèves atteignent ce seuil indépendamment du niveau scolaire. Les différences observées, quoique significatives, sont néanmoins relativement faibles considérant qu'elles semblent plafonner autour du seuil minimal de réalisme.

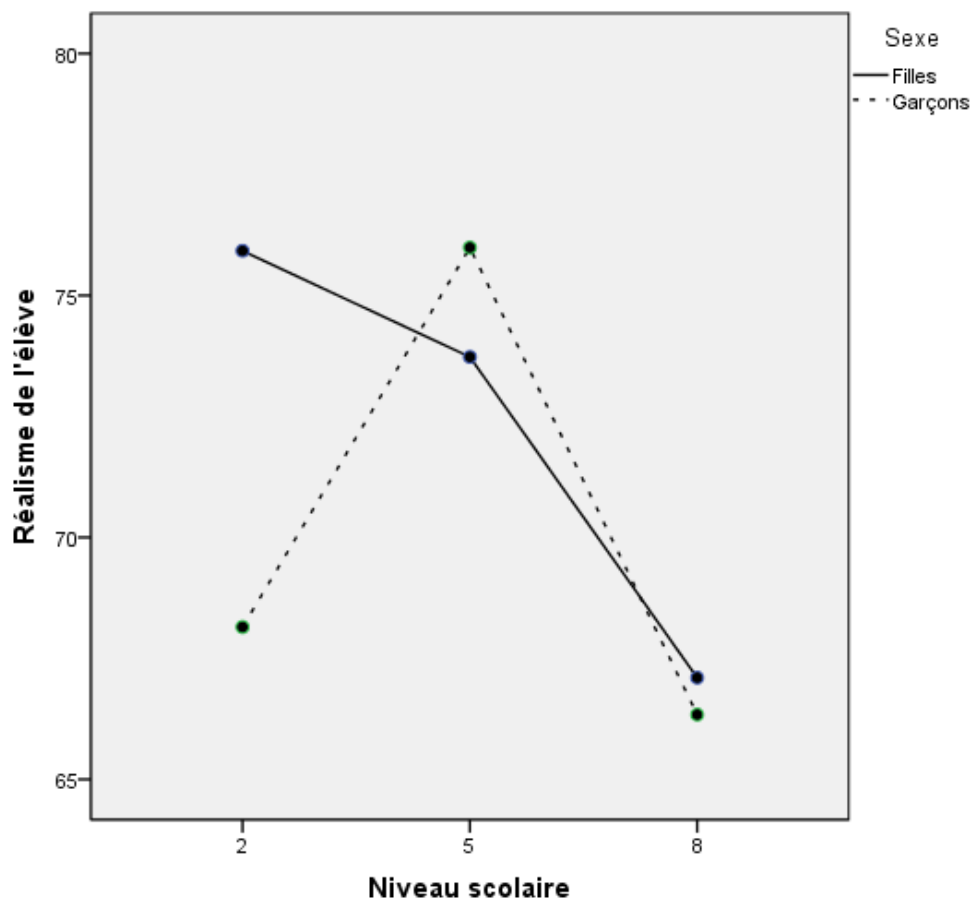


Figure 23. Réalisme moyen des élèves selon le niveau scolaire et le sexe pour l'épreuve des Figures graduées.

### Épreuve des Concentrations

Le Tableau 22 a présenté les résultats de l'analyse univariée de la variance pour les deux épreuves. Dans le cas de l'épreuve des Concentrations, nous constatons une différence significative pour le sexe ( $F = 8,523$  ;  $p = 0,004$ ) et le niveau scolaire ( $F = 171,455$  ;  $p < 0,000$ ). Il faut aussi noter l'absence d'interaction significative entre le sexe et le niveau scolaire ( $F =$

3,000 ;  $p = 0,051$ ). Notons que le devis ainsi que le nombre d'élèves nous ont permis d'avoir suffisamment de puissance avec respectivement 83,0 %, 100 % et 58,2 % de probabilité de rejeter les hypothèses nulles lorsqu'elles étaient fausses. Nous notons des valeurs d'êta-carré assez faibles pour le sexe (0,014) et l'interaction (0,010), ce qui indique qu'une très faible proportion de la variance est expliquée par les différents niveaux de ces conditions. La valeur d'êta-carré pour le niveau scolaire est plus élevée (0,366). Le  $R$  carré global indique que 36,8 % de la variance du réalisme est expliquée par les variables indépendantes. Notons que, contrairement à l'épreuve des Figures graduées, le test d'homogénéité de la variance révèle que la variance des erreurs du réalisme des élèves est homogène entre les différents groupes ( $F = 1,168$  ;  $p = 0,324$ ).

L'analyse univariée de la variance a permis de constater qu'il y a une différence significative entre les moyennes des filles et des garçons. Tous les tests de comparaison effectués *a posteriori* montrent des différences significatives sans équivoque entre les élèves de 2<sup>e</sup> année et les élèves de 5<sup>e</sup> année, entre les élèves de 5<sup>e</sup> année et les élèves de 8<sup>e</sup> année et entre les élèves de 2<sup>e</sup> année et les élèves de 8<sup>e</sup> année. À la lumière de ces résultats, nous pouvons rejeter l'hypothèse nulle concernant l'effet principal lié au sexe de l'élève. Nous rejetons également l'hypothèse nulle concernant l'effet principal lié aux niveaux scolaires ou à l'âge de l'élève. Les résultats soutiennent que le réalisme des élèves varie en fonction de son niveau scolaire.

La Figure 24 permet de mieux visualiser la relation entre le réalisme des élèves et le niveau scolaire. Nous constatons aussi des différences de réalisme entre les garçons et les filles ainsi qu'entre les niveaux scolaires. Nous constatons cependant une tendance haussière et quasi linéaire avec le niveau scolaire. Notons que la différence de réalisme entre les garçons et les filles

s'observe surtout chez les élèves de deuxième et cinquième année, et ce, dans le cas des deux épreuves. Il apparaît que chez les élèves moins réalistes, les garçons tendent à être moins réalistes que les filles.

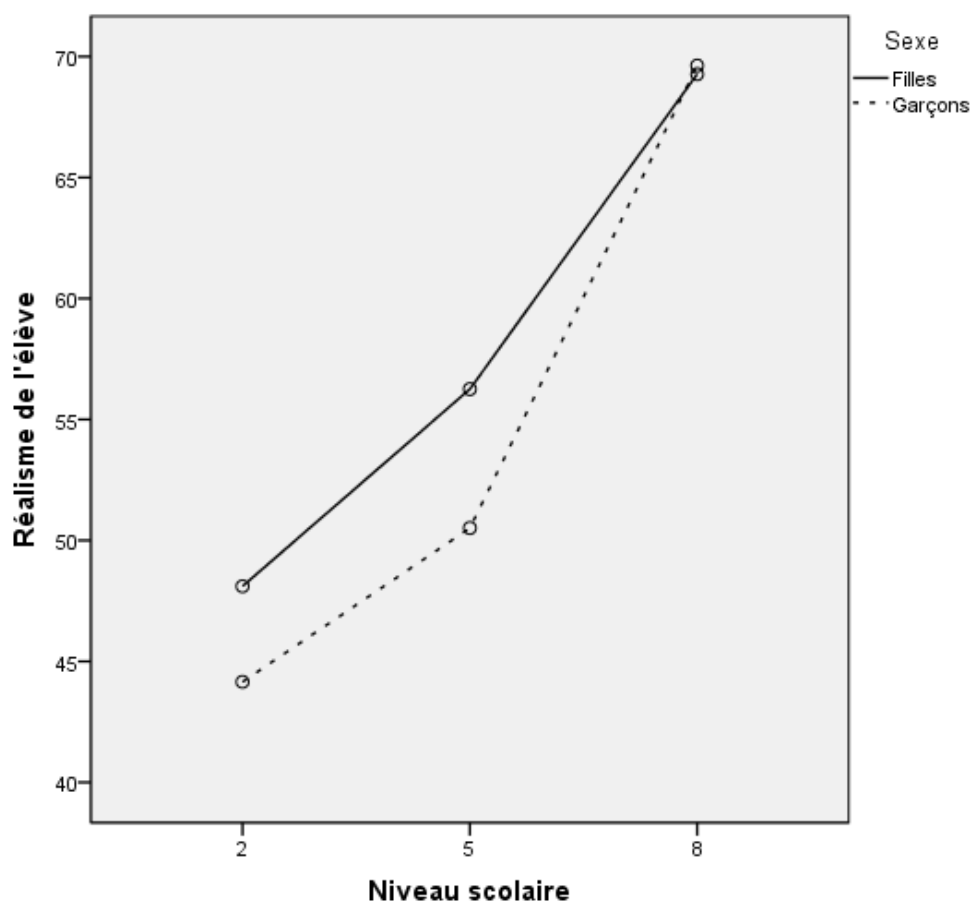


Figure 24. Réalisme moyen des élèves selon le niveau scolaire et le sexe pour l'épreuve des Concentrations.

### **Réalisme des élèves et domaine cognitif : comparaison entre les garçons et les filles**

Le Tableau 24 présente les résultats de l'analyse de la variance selon un plan factoriel à trois critères de classification (sexe x niveau scolaire x épreuve). Le réalisme moyen est différent d'une épreuve à l'autre ( $F = 324,715$  ;  $p < 0,000$ ). Nous constatons une différence significative du réalisme moyen en fonction du niveau scolaire ( $F = 91,282$  ;  $p < 0,000$ ) ce qui suggère que le réalisme moyen varie significativement en fonction du niveau scolaire. Il y a une interaction significative entre le sexe et le niveau scolaire ( $F = 3,049$  ;  $p = 0,048$ ). Il n'y a pas de différence significative entre les garçons et les filles. Dans les cas de différences significatives, notons que le devis ainsi que le nombre d'élèves ont permis d'avoir suffisamment de puissance avec respectivement 100 %, 100 % et 59,0 % de probabilité de rejeter les hypothèses nulles lorsqu'elles étaient fausses. Ceci n'est pas le cas pour le sexe avec une puissance de 0,065. Les valeurs d'êta-carré sont nulles pour le sexe (0,000) et l'interaction entre le sexe et le niveau scolaire (0,005), ce qui indique qu'une faible proportion de la variance est expliquée par les différents niveaux de ces conditions. Les valeurs d'êta-carré sont plus importantes pour le niveau scolaire (0,123) et l'épreuve (0,199) ce qui indique qu'une assez bonne proportion de la variance est expliquée par les différents niveaux scolaires et le domaine cognitif de l'épreuve. Le  $R$  carré global avec 32,4 % de la variance du réalisme expliquée par les variables indépendantes combinées suggère que les liens entre les variables indépendantes et la variable dépendante sont assez forts. Le Tableau 25 montre que l'interaction est significative tant entre les élèves de la 2<sup>e</sup> année à la 8<sup>e</sup> année ( $p < 0,000$ ) que les élèves de la 5<sup>e</sup> année à la 8<sup>e</sup> année ( $p < 0,000$ ). Le test d'homogénéité de la variance de Levene ne révèle pas d'inégalités entre les différents groupes ( $F = 1,578$  ;  $p = 0,099$ ).

Tableau 24

*Tableau résumé de l'analyse univariée de la variance selon un plan à trois critères de classification*

Source	SC	DI	CM	F	Sig.	Êta-carré partiel	Puissance
Domaine de l'épreuve	97590,788	1	97590,788	324,715	,000	,199	1,000
Sexe	38,975	1	38,975	,130	,719	,000	,065
Niveau scolaire	54868,285	2	27434,142	91,282	,000	,123	1,000
Sexe * niveau scolaire	1832,861	2	916,430	3,049	,048	,005	,590
Erreur	391908,398	1304	300,543				
Alpha = 0,05	R carré = 0,324						

Note : Les variables indépendantes sont des variables catégorielles traitées à titre de facteurs fixes. La variable dépendante est l'indice de réalisme.

Tableau 25

*Résultats des contrastes pour les deux épreuves en fonction du niveau scolaire*

Niveau 1 et niveau 3	Estimation de contraste	-6,203
	Différence (estimation - hypothèse)	-6,203
	Erreur	1,177
	Sig.	,000
Niveau 2 et niveau 3	Estimation de contraste	-4,188
	Différence (estimation - hypothèse)	-4,188
	Erreur	1,199
	Sig.	,000

Légende : Niveau 1, 2 et 3 réfèrent aux niveaux scolaires correspondant à la 2<sup>e</sup> année, 5<sup>e</sup> année et 8<sup>e</sup> année respectivement. Il s'agit de contrastes simples qui visent à montrer les différences entre les niveaux.

L'analyse de la variance univariée en plan factoriel a permis de constater qu'il y a une différence significative du réalisme moyen des élèves selon le domaine cognitif de l'épreuve. À la lumière de ces résultats, nous pouvons rejeter l'hypothèse nulle concernant l'effet principal lié au domaine cognitif de l'épreuve. Les résultats soutiennent que le réalisme de l'élève varie en fonction de son niveau scolaire et non en fonction du sexe. L'interaction entre le domaine cognitif de l'épreuve et le niveau scolaire est significative. Les Figures 25 et 26 permettent de mieux visualiser la relation entre le réalisme moyen des élèves dans chacune des épreuves et leur niveau scolaire. Il semble que le réalisme varie considérablement en fonction du domaine cognitif en 2<sup>e</sup> et en 5<sup>e</sup> année. Cette différence s'estompe en 8<sup>e</sup> année. Il faut dire que dans le cas de l'épreuve des Figures graduées, le réalisme moyen des élèves plafonne autour de 0,71. Le réalisme moyen des élèves pour l'épreuve des Concentrations est plus faible et rattrape le réalisme moyen pour l'épreuve des Figures graduées à mesure que le niveau scolaire augmente.

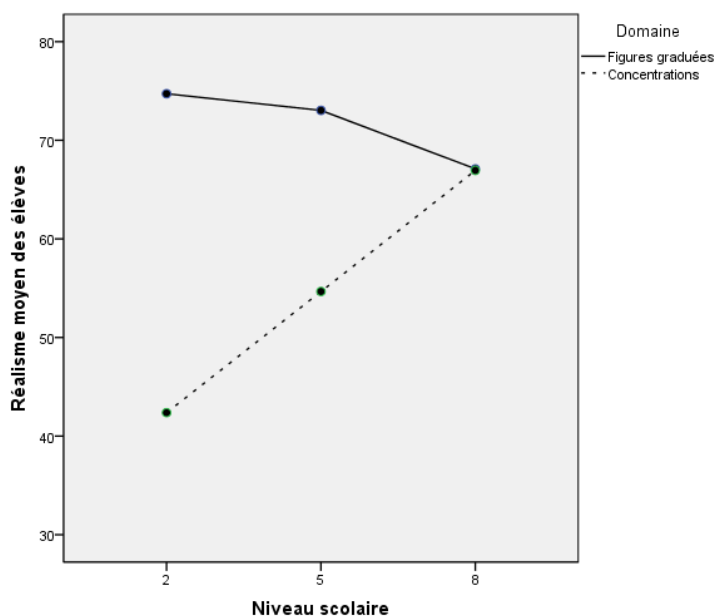


Figure 25. Réalisme moyen des filles dans les deux épreuves selon le niveau scolaire.

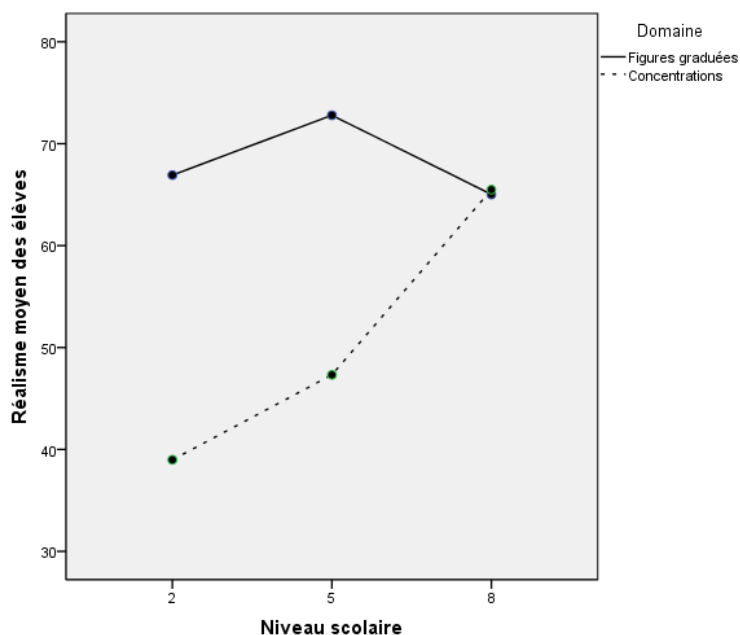


Figure 26. Réalisme moyen des garçons dans les deux épreuves selon le niveau scolaire.

Le réalisme des élèves est donc dépendant du domaine cognitif de la tâche. Notons que cette différence pourrait être attribuable aux modalités de *testing* des deux épreuves. L'épreuve des Figures graduées permet à l'élève de constater visuellement en temps réel la qualité de reproduction de ses dessins, ce qui n'est pas possible dans le cas de l'épreuve des Concentrations portant sur le choix de la proportion ou du ratio le plus élevé.

### Réalisme selon l'écart entre l'exigence cognitive de la tâche et le stade atteint

Nous faisons l'hypothèse que l'élève est plus réaliste lorsque le niveau cognitif atteint par celui-ci est supérieur ou égal à l'exigence cognitive de la tâche. L'élève est moins réaliste si son niveau de développement est inférieur à l'exigence cognitive de la tâche. Rappelons que nous faisons l'hypothèse que la diminution de réalisme suit un modèle curvilinéaire tel que présenté

antérieurement à la Figure 4. En d'autres mots, le réalisme de l'élève est plus élevé dans les tâches qui présentent des exigences cognitives égales ou inférieures au stade cognitif atteint par ce dernier.

Nous avons comparé le réalisme des élèves dans les tâches qui correspondent à leur niveau opératoire ou qui leur sont inférieures avec le réalisme manifesté dans les tâches qui sont de niveau opératoire supérieur. Par exemple, pour un élève ayant atteint le stade cognitif IIA dans l'épreuve des Concentrations, nous comparons la moyenne du réalisme de cet élève dans les tâches de niveaux IA, IB, IIA avec la moyenne de son réalisme dans les tâches de niveaux IIB, IIIA, IIIA2, IIIB.

Nous avons examiné la relation entre le sexe et le réalisme des élèves (*Rs*) dans le cas des items de niveau d'exigence égale ou inférieure au niveau de développement de l'élève avec les items d'exigence supérieure. En fait, nous avons calculé le réalisme des élèves pour les items de niveau d'exigence égale ou inférieure au niveau de développement de l'élève et nous avons fait le même calcul pour les items d'exigence supérieure. Nous avons contrôlé cette variable à titre de variable indépendante. Une analyse multivariée de la variance a permis d'en apprendre plus sur ces liens et sur les interactions potentielles entre les variables indépendantes. Les Tableaux 26 et 27 qui suivent présentent certaines statistiques descriptives sur le réalisme de l'élève utilisé à titre de variable dépendante pour l'analyse multivariée de la variance. Les statistiques présentées sont en l'occurrence la moyenne, l'écart-type et la variance du réalisme en fonction des variables indépendantes telles le sexe et le niveau scolaire de l'élève. Nous notons que la variance du réalisme des élèves semble se distribuer inégalement. Pour les tâches d'exigence cognitive supérieure au niveau de développement cognitif atteint par les élèves, l'écart-type et la variance

sont moins élevés. Ceci suggère que cette distribution tend à présenter une voussure qui a une forme plus leptokurtique que pour les autres élèves. Il y aurait donc moins d'écarts de réalisme entre ces élèves.

Tableau 26

*Statistiques descriptives sur le réalisme selon l'écart entre l'exigence cognitive de la tâche et le stade atteint - Épreuve des Figures graduées*

Exigence	Sexe	Moyenne	Écart-type	Variance
Niveau atteint	F	61,76	28,90	834,93
	G	61,74	28,37	804,82
Supérieur	F	57,95	25,14	632,24
	G	57,41	24,51	600,54
Total	F	59,85	27,13	736,16
	G	59,57	26,57	706,22

Tableau 27

*Statistiques descriptives sur le réalisme selon l'écart entre l'exigence cognitive de la tâche et le stade atteint - Épreuve des Concentrations*

Sexe	Niveau	Moyenne	Écart-type	Variance
Niveau atteint	F	89,934	11,485	131,905
	G	89,873	14,066	197,860
Supérieur	F	34.035	14.018	196.513
	G	29.009	13.927	193.968
Total	F	61.985	30.762	946.307
	G	59.441	33.515	1123.281

Le Tableau 28 présente les résultats de l'analyse de la variance à mesures répétées sur les écarts de réalisme entre l'exigence cognitive de la tâche et le stade atteint pour l'épreuve des Figures graduées et pour l'épreuve des Concentrations. Nous constatons que les deux effets principaux sont significatifs. En utilisant le Lambda de Wilks à titre de critère, nous constatons une différence significative du réalisme moyen pour les items d'exigence cognitive au niveau atteint ou inférieur et celles d'exigences supérieures ( $F = 2966,063$  ;  $p < 0,000$ ) et entre les filles et les garçons ( $F = 5,634$  ;  $p = 0,004$ ). Il faut aussi noter une interaction significative entre le sexe et l'écart entre le réalisme pour les items d'exigence cognitive au niveau atteint ou inférieur et celles d'exigences supérieures ( $F = 5,376$  ;  $p = 0,005$ ). Notons que le devis ainsi que le nombre d'élèves ont permis d'avoir suffisamment de puissance avec respectivement 100 %, 86,1 % et 84,3 % de probabilité de rejeter les hypothèses nulles lorsqu'elles étaient fausses. Les valeurs d'êta-carré sont assez faibles pour le sexe (0,009) et l'interaction (0,009), ce qui indique qu'une faible proportion de la variance est expliquée par les différents niveaux de ces conditions. La valeur d'êta-carré pour l'écart entre le réalisme pour les items d'exigence cognitive au niveau atteint ou inférieur et celles d'exigences supérieures (0,827) est très importante. Puisque les liens sont significatifs, des analyses univariées sont recommandées et sont présentées dans la prochaine section. Le test de Box confirme que l'homogénéité de la variance n'est pas garantie ( $F = 5,165$  ;  $p < 0,000$ ).

Tableau 28

*Tableau résumé de l'analyse de la variance à mesures répétées sur les écarts de réalisme entre l'exigence cognitive de la tâche et le stade atteint pour l'épreuve des Figures graduées et pour l'épreuve des Concentrations*

Source		Valeur	<i>DI</i>	<i>F</i>	Sig.	Êta-carré partiel	Puissance
Écart	Pillai	,827	2,000	2966,023 <sup>a</sup>	0,000	,827	1,000
	Wilks	,173	2,000	2966,023 <sup>a</sup>	0,000	,827	1,000
	Hotelling	4,765	2,000	2966,023 <sup>a</sup>	0,000	,827	1,000
	Roy	4,765	2,000	2966,023 <sup>a</sup>	0,000	,827	1,000
Sexe	Pillai	,009	2,000	5,634 <sup>a</sup>	0,004	,009	,861
	Wilks	,991	2,000	5,634 <sup>a</sup>	0,004	,009	,861
	Hotelling	,009	2,000	5,634 <sup>a</sup>	0,004	,009	,861
	Roy	,009	2,000	5,634 <sup>a</sup>	0,004	,009	,861
Écart * sexe	Pillai	,009	2,000	5,376 <sup>a</sup>	0,005	,009	,843
	Wilks	,991	2,000	5,376 <sup>a</sup>	0,005	,009	,843
	Hotelling	,009	2,000	5,376 <sup>a</sup>	0,005	,009	,843
	Roy	,009	2,000	5,376 <sup>a</sup>	0,005	,009	,843

Note : Les variables indépendantes sont des variables catégorielles traitées à titre de facteurs fixes. Les variables dépendantes sont l'indice de réalisme pour chaque épreuve. La variable Écart a deux niveaux, soit inférieur et supérieur. Nous avons calculé le réalisme des élèves pour les items de niveau d'exigence égale ou inférieure au niveau de développement de l'élève et nous avons fait le même calcul pour les items d'exigence supérieure.

Le Tableau 29 présente les résultats de l'analyse de la variance à mesures répétées sur les écarts de réalisme entre l'exigence cognitive de la tâche et le stade atteint pour l'épreuve des Figures graduées et l'épreuve des Concentrations. Nous constatons une différence significative

du réalisme moyen pour les items d'exigence cognitive au niveau atteint ou inférieur et ceux d'exigences supérieures ( $F = 6,493$  ;  $p = 0,011$ ) ce qui suggère que le réalisme moyen varie significativement en fonction de l'écart entre le degré d'exigence cognitive de la tâche au niveau cognitif atteint par l'élève et le degré d'exigence cognitive de la tâche qui excède le niveau cognitif atteint par l'élève. Il n'a pas de différence entre les filles et les garçons ( $F = 0,016$  ;  $p = 0,901$ ). Il n'y a pas non plus d'interaction significative entre l'écart d'exigence cognitive et le sexe. Dans les cas de différences significatives, notons que le devis ainsi que le nombre d'élèves ont permis d'avoir suffisamment de puissance avec 72,1 % de probabilité de rejeter les hypothèses nulles lorsqu'elles étaient fausses. Ceci n'est pas le cas pour le sexe avec une puissance de 0,052. Les valeurs d'éta-carré sont très faibles pour le sexe (0,000) et l'interaction entre le sexe et l'écart de réalisme (0,000), ce qui indique qu'une faible proportion de la variance est expliquée par les différents niveaux de ces conditions. Le  $R$  carré de 0,5 % confirme également la faible proportion de la variance expliquée. Le test d'homogénéité de la variance révèle un manque de conformité au postulat ( $F = 5,021$  ;  $p = 0,002$ ).

L'analyse de la variance multivariée a permis de constater que le réalisme moyen varie significativement en fonction de l'écart entre le degré d'exigence cognitive de la tâche au niveau cognitif atteint par l'élève et le degré d'exigence cognitive de la tâche qui excède le niveau cognitif atteint par l'élève. À la lumière de ces résultats, nous pouvons rejeter l'hypothèse nulle concernant l'effet principal lié à l'écart d'exigence cognitive. Nous notons néanmoins que les analyses sont affaiblies par le manque de conformité au postulat d'homoscédasticité. La Figure 27 permet de mieux visualiser la relation entre le réalisme moyen des élèves en fonction

de l'écart entre le degré d'exigence cognitive de la tâche au niveau cognitif atteint par l'élève et le degré d'exigence cognitive de la tâche qui excède le niveau cognitif atteint par l'élève.

Tableau 29

*Tableau résumé de l'analyse de la variance à mesures répétées sur les écarts de réalisme entre l'exigence cognitive de la tâche et le stade atteint pour l'épreuve des Figures graduées et pour l'épreuve des Concentrations traitées séparément*

Source		SC	Dl	CM	F	Sig.	Êta-carré	
							partiel	Puissance
Écarts de réalisme	FG	4740,193	1	4740,193	6,493	0,011	,005	,721
	CONC	1060542,611	1	1060542,611	5924,208	0,000	,826	1,000
Sexe	FG	11,405	1	11,405	0,016	0,901	,000	,052
	CONC	2013,583	1	2013,583	11,248	0,001	,009	,918
Écarts de réalisme * sexe	FG	22,018	1	22,018	0,030	0,862	,000	,053
	CONC	1917,912	1	1917,912	10,713	0,001	,009	,905
Erreur	FG	909673,581	1246	730,075				
	CONC	223057,010	1246	179,018				
Total	FG	5297184,950	1250					
	CONC	5906402,726	1250					

Alpha = 0,05

R carré = 0,005 pour FG et 0,827 pour CONC

Note : Les variables indépendantes sont des variables catégorielles traitées à titre de facteurs fixes. Les variables dépendantes sont l'écart de l'indice de réalisme pour chaque épreuve, soit FG pour Figures graduées et CONC pour Concentrations.

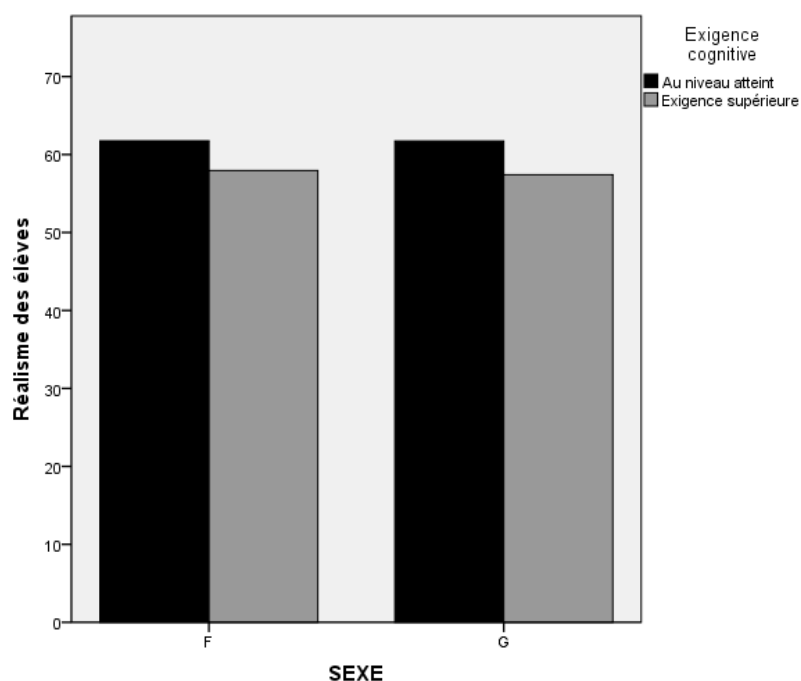


Figure 27. Écart du réalisme moyen des élèves de tous les niveaux scolaires et l'exigence cognitive de la tâche pour l'épreuve des Figures graduées.

Comme dans le cas de l'épreuve des Figures graduées, le Tableau 29 présente également les résultats de l'analyse de la variance à mesures répétées sur les écarts de réalisme entre l'exigence cognitive de la tâche et le stade atteint pour l'épreuve des Concentrations. Nous constatons une différence significative du réalisme moyen pour les items d'exigence cognitive au niveau atteint ou inférieur et ceux d'exigences supérieures ( $F = 5924,208$  ;  $p < 0,000$ ) ce qui suggère que le réalisme moyen varie significativement en fonction de l'écart entre le degré d'exigence cognitive de la tâche au niveau cognitif atteint par l'élève et le degré d'exigence cognitive de la tâche qui excède le niveau cognitif atteint par l'élève. Nous notons une différence entre les filles et les garçons ( $F = 11,248$  ;  $p = 0,001$ ). Il y a une interaction significative entre

l'écart d'exigence cognitive et sexe de l'élève ( $F = 10,713$  ;  $p = 0,001$ ). Dans les cas de différences significatives, notons que le devis ainsi que le nombre d'élèves ont permis d'avoir suffisamment de puissance avec respectivement 100 %, 91,8 % et 90,5 % de probabilité de rejeter les hypothèses nulles lorsqu'elles étaient fausses. Les valeurs d'éta-carré sont assez faibles pour l'interaction avec le sexe (0,009) ce qui indique qu'une faible proportion de la variance est expliquée par les différents niveaux de ces conditions. La valeur d'éta-carré est forte pour l'écart d'exigence cognitive (0,826) ce qui indique qu'une forte proportion de la variance est expliquée par les différents niveaux de ces conditions. Le  $R$  carré de 82,7% confirme également ce constat. Le test d'homogénéité de la variance révèle un manque de conformité au postulat pour le réalisme ( $F = 6,200$  ;  $p < 0,000$ ).

L'analyse de la variance multivariée a permis de constater que le réalisme moyen varie significativement en fonction de l'écart entre le degré d'exigence cognitive de la tâche au niveau cognitif atteint par l'élève et le degré d'exigence cognitive de la tâche qui excède le niveau cognitif atteint par l'élève. À la lumière de ces résultats, nous pouvons rejeter l'hypothèse nulle concernant l'effet principal lié à l'écart d'exigence cognitive. Les résultats soutiennent que le réalisme de l'élève varie en fonction du sexe de l'élève. Nous notons néanmoins que les analyses sont affaiblies par le manque de conformité de certains postulats. La Figure 28 permet de mieux visualiser la relation entre le réalisme moyen des élèves et le sexe.

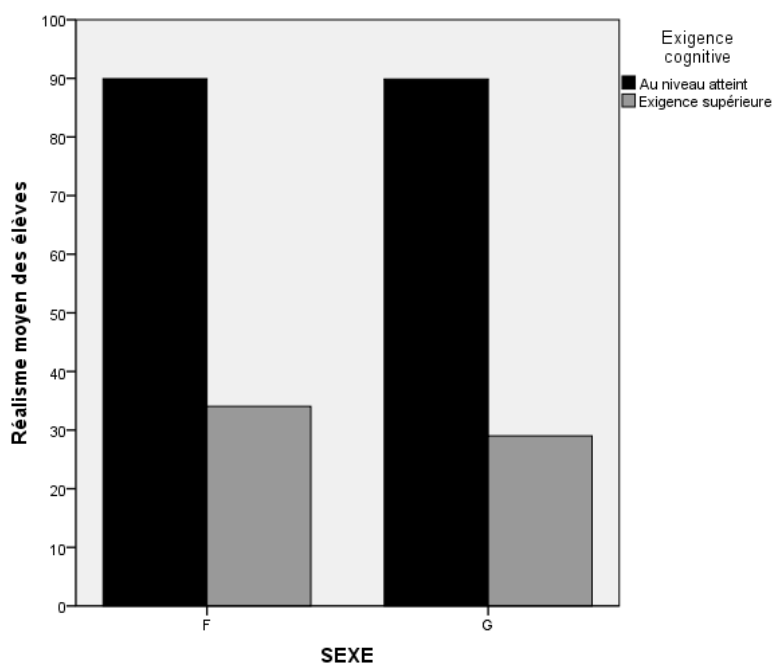


Figure 28. Écart du réalisme moyen des élèves selon le sexe de l'élève et l'exigence cognitive de la tâche pour l'épreuve des Concentrations.

### **Le réalisme de l'élève dans les cas extrêmes d'écart entre l'exigence cognitive de la tâche et le niveau de développement cognitif atteint**

Antérieurement, nous avons suggéré que dans les cas extrêmes d'items faciles et difficiles, l'élève pourrait manifester un haut degré de réalisme puisqu'il serait certain de savoir ou de ne pas savoir, et ce, avec raison. On proposait, tel que montré dans la Figure 4, une relation en forme de « U ». Les Tableaux 30 et 31 montrent le réalisme moyen des élèves pour les deux épreuves à mesure que l'exigence cognitive de la tâche s'éloigne du niveau de développement cognitif atteint par l'élève. Nous nous sommes limités aux niveaux allant de IA à IIA. Les élèves des niveaux O et VI sont peu nombreux et plusieurs d'entre eux présentent des difficultés d'apprentissage ou ont des besoins particuliers. Nous les excluons de cette analyse puisqu'il y a

probablement des facteurs ou des variables atténuatrices qui ne sont pas contrôlés. Nous excluons également les élèves des niveaux supérieurs à IIA puisqu'il n'y a pas d'écart suffisant entre le niveau de développement atteint par les élèves et l'exigence cognitive de la tâche pour constater la linéarité ou l'absence de linéarité de la diminution de réalisme.

À l'aide des Figures 29 à 35, nous observons que le réalisme moyen des élèves diminue progressivement à mesure que l'exigence cognitive de la tâche dépasse leur stade atteint. Le réalisme tend à remonter lorsque l'exigence cognitive de la tâche s'éloigne considérablement du stade cognitif atteint par l'élève. Par exemple, la Figure 29 présente la diminution progressive du réalisme des élèves pour les items qui excèdent le niveau de développement cognitif de l'élève situé à IA. Une analyse de tendance montre que, sauf dans un seul cas, les courbes sont curvilinéaires et en forme de U. Les coefficients de détermination  $R$  carré des fonctions quadratiques sont tous supérieurs à 0,86. Notons qu'à la Figure 31, le coefficient de détermination est forcément à 1,000 puisqu'il n'y a que 3 points. Ce constat soutient notre proposition que la diminution de réalisme pour les tâches qui excède le niveau cognitif atteint par les élèves n'est pas linéaire et probablement en forme de « U » tel que proposé.

Tableau 30

*Réalisme moyen des élèves à mesure que l'exigence cognitive de la tâche s'éloigne du niveau cognitif atteint par l'élève – épreuve des Figures graduées*

		Exigence cognitive de la tâche						
Stade atteint								
par l'élève	n	OA	OB	IA	IB	IIA	IIB	IIIA
IA	101	74,2	68,0	75,1	60,3	53,0	47,3	60,2
IB	114	74,6	73,5	75,2	69,9	53,8	45,6	60,3
IIA	195	69,8	72,6	74,0	63,2	67,4	54,7	64,2

Tableau 31

*Réalisme moyen des élèves à mesure que l'exigence cognitive de la tâche s'éloigne du niveau cognitif atteint par l'élève – épreuve des Concentrations*

		Exigence cognitive de la tâche							
Stade atteint									
par l'élève	n	IA	IB	IC	IIA	IIB	IIIA1	IIIA2	IIIB
IA	132	94,0	13,0	10,8	8,5	8,0	52,5	52,8	49,7
IB	139	94,9	89,0	32,6	17,1	14,1	46,8	44,3	41,3
IC	153	95,9	89,3	81,9	29,8	19,0	49,1	46,1	37,1
IIA	123	96,0	90,6	82,0	84,6	30,8	30,6	28,9	25,5

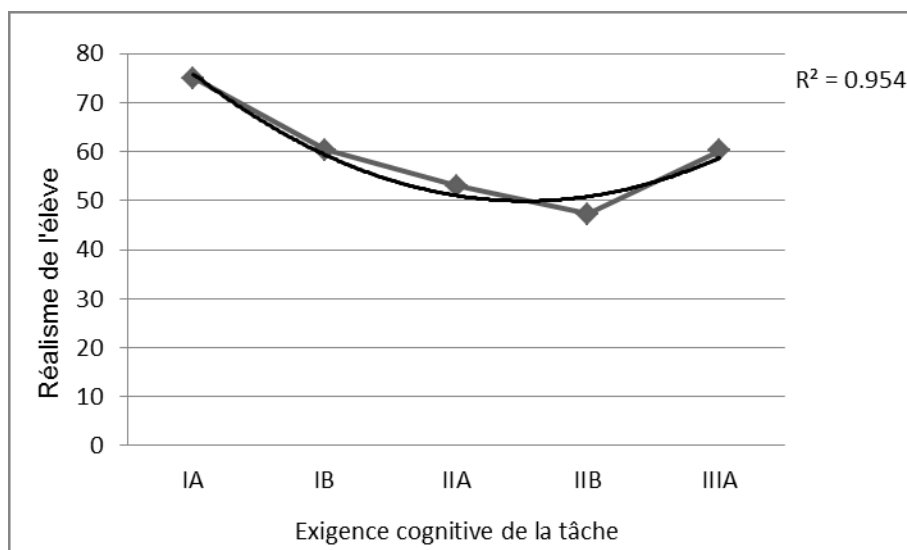


Figure 29. Réalisme moyen des élèves du niveau IA à mesure que l'exigence cognitive de la tâche s'éloigne du niveau cognitif atteint par l'élève – épreuve des Figures graduées.

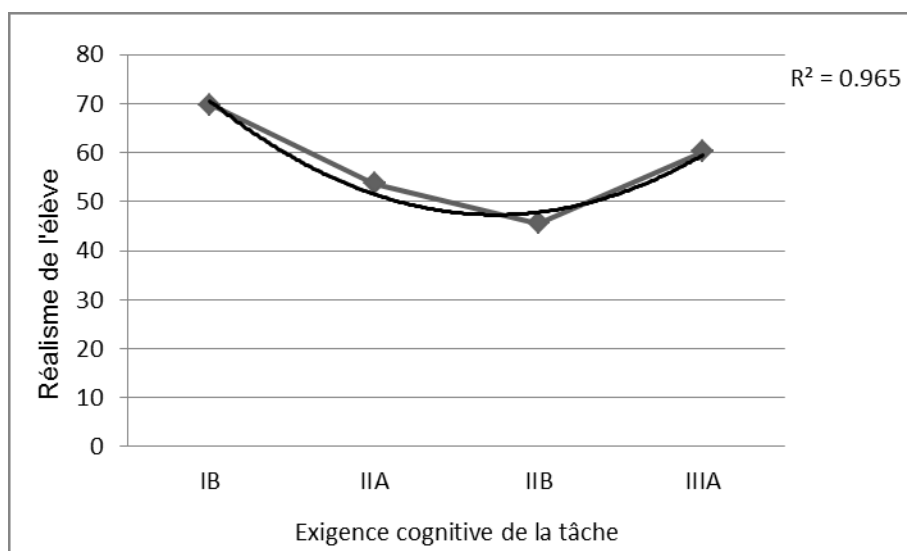


Figure 30. Réalisme moyen des élèves du niveau IB à mesure que l'exigence cognitive de la tâche s'éloigne du niveau cognitif atteint par l'élève – épreuve des Figures graduées.

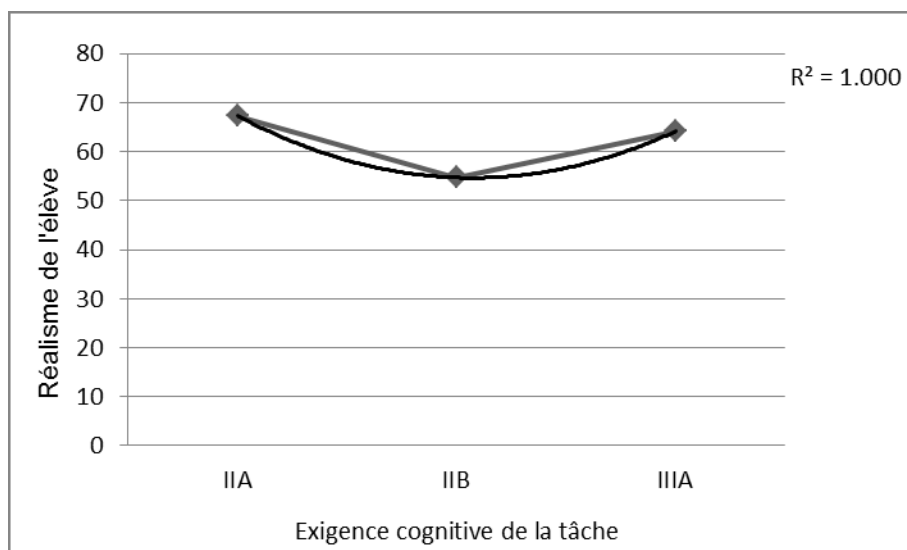


Figure 31. Réalisme moyen des élèves du niveau IIA à mesure que l'exigence cognitive de la tâche s'éloigne du niveau cognitif atteint par l'élève – épreuve des Figures graduées.

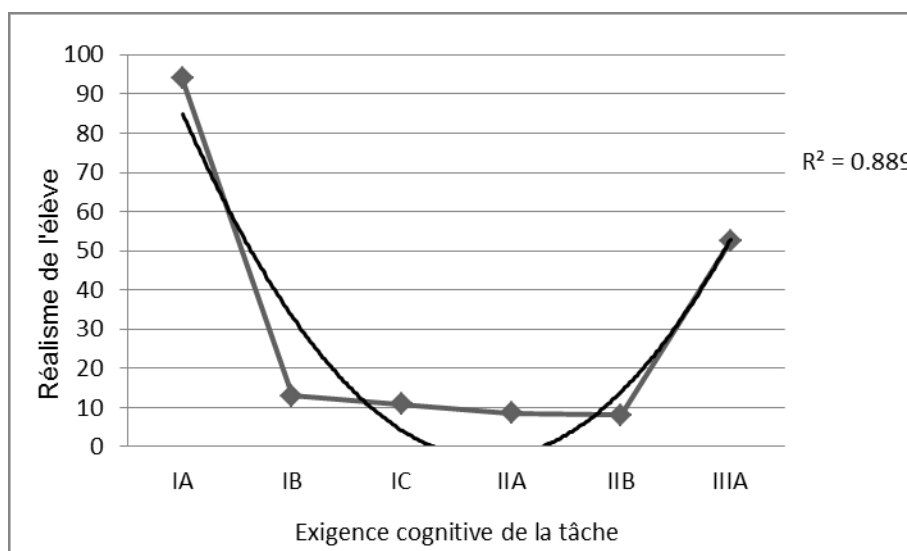


Figure 32. Réalisme moyen des élèves du niveau IA à mesure que l'exigence cognitive de la tâche s'éloigne du niveau cognitif atteint par l'élève – épreuve des Concentrations.

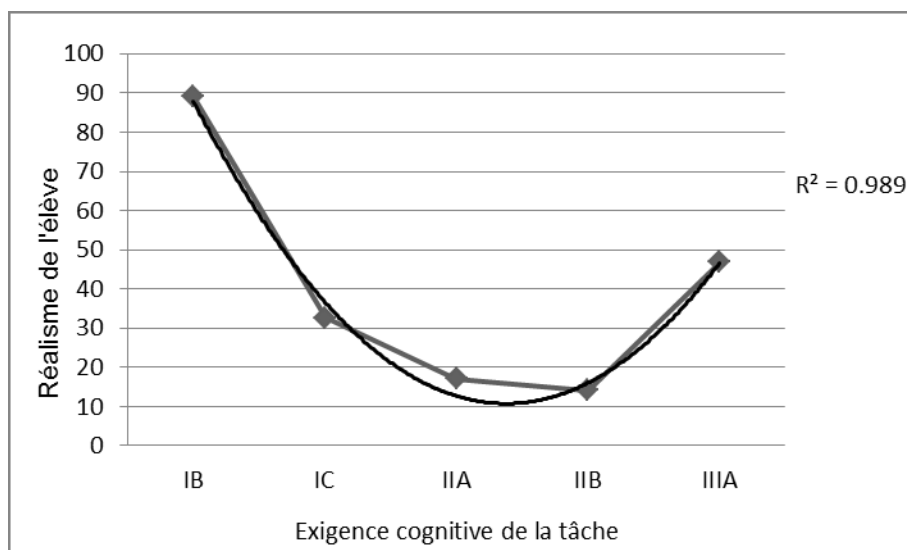


Figure 33. Réalisme moyen des élèves du niveau IB à mesure que l'exigence cognitive de la tâche s'éloigne du niveau cognitif atteint par l'élève – épreuve des Concentrations.

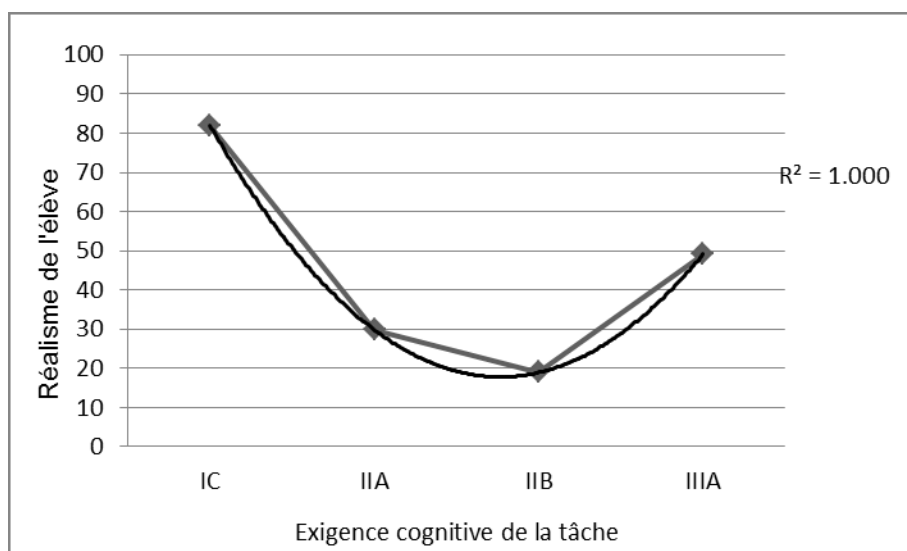
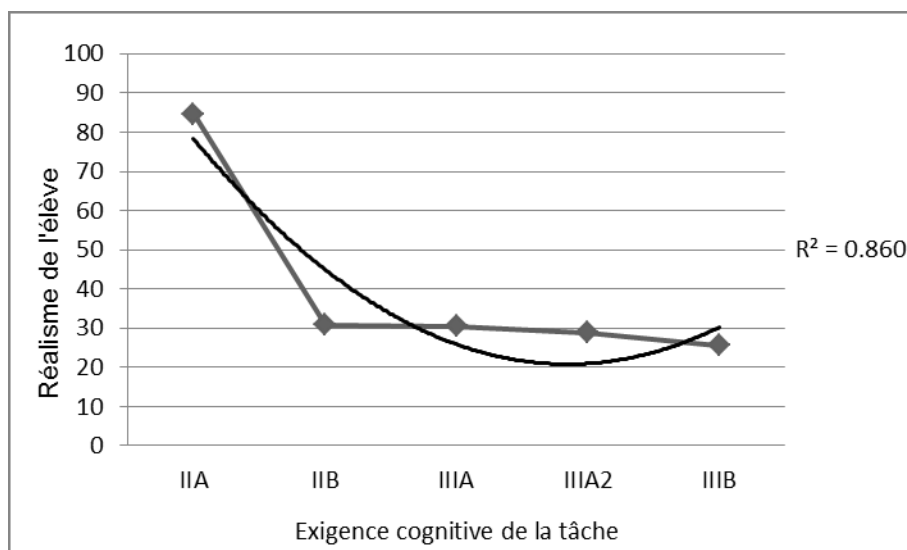


Figure 34. Réalisme moyen des élèves du niveau IC à mesure que l'exigence cognitive de la tâche s'éloigne du niveau cognitif atteint par l'élève – épreuve des Concentrations.



*Figure 35.* Réalisme moyen des élèves du niveau IIA à mesure que l'exigence cognitive de la tâche s'éloigne du niveau cognitif atteint par l'élève – épreuve des Concentrations.

## DISCUSSION

Cette recherche proposait de vérifier si l'évaluation du réalisme a un intérêt diagnostique et peut donner lieu à des résultats valides. Pour ce faire, il faut connaître les conditions entourant son utilisation comme le domaine cognitif de la tâche, l'influence des différences individuelles comme le sexe de l'élève, son stade opératoire atteint ou son niveau scolaire. Nous cherchons entre autres à déterminer l'âge ou le stade de développement minimal requis chez la majorité des élèves pour une utilisation valide du réalisme dans un contexte scolaire. Également, nous avons fait l'hypothèse que le réalisme des élèves pourrait varier en fonction de l'écart entre la complexité de la tâche et leurs capacités cognitives selon leur stade de développement. Dans cette section, nous allons jeter un dernier regard sur les résultats obtenus et sur les pistes de recherche à considérer.

La recension des écrits nous permettait d'anticiper que le réalisme pouvait varier en fonction de certaines différences individuelles comme l'âge. Il était ainsi raisonnable de croire qu'un élève plus développé sur le plan cognitif est plus réaliste dans son appréciation de la qualité de ses réponses. Nous pouvions aussi présumer que les items difficiles ou qui font appel à des habiletés supérieures de la pensée entraînent une baisse de réalisme surtout lorsque l'exigence cognitive de la tâche dépasse le niveau cognitif atteint par l'élève. La nature de la tâche en matière de domaine cognitif ou de sujet pourrait avoir un impact sur le réalisme.

Les premières analyses consistaient à vérifier les liens entre les variables de cette étude. Des analyses de régression multiple linéaire pour les deux épreuves ont été effectuées. Les indices de multicolinéarité ont été vérifiés et on constate que les analyses sont conformes à cet

égard. Le coefficient de détermination de 0,45 % pour l'épreuve des Figures graduées indique que la variance du réalisme expliquée par les variables indépendantes est très faible. Pour l'épreuve des Concentrations, 57,6 % de la variance du réalisme est expliquée par le stade cognitif atteint, le niveau scolaire et le sexe de l'élève. La faiblesse du devis constaté par l'effet de plafonnement explique en partie la faiblesse des liens entre les variables indépendantes et le réalisme en ce qui concerne l'épreuve des Figures graduées.

Nous remarquons que les coefficients de corrélation partielle des variables indépendantes sont faibles. Ils sont plus élevés pour l'épreuve des Concentrations que pour l'épreuve des Figures graduées. Le stade cognitif atteint et le niveau scolaire sont plus corrélés avec le réalisme de l'élève que le sexe.

Une seconde analyse portait sur les corrélations non paramétriques entre le réalisme des élèves et les stades atteints dans les deux épreuves. Dans le cas de l'épreuve des Concentrations, les stades atteints et le réalisme des élèves sont corrélés significativement. La corrélation entre ces mêmes variables n'est pas significative pour l'épreuve des Figures graduées. On constate qu'il y a un lien entre les stades atteints tels que déterminés par les deux épreuves, ce qui était attendu. On conclut que les deux épreuves permettent de déterminer le stade atteint.

Nous constatons un effet de plafonnement pour l'épreuve des Figures graduées autour du niveau de réalisme correspondant à 0,71. Nous notons aussi que ce phénomène s'explique en partie en observant le réalisme moyen des élèves de 8<sup>e</sup> année. Nous avons postulé que le réalisme moyen de ces derniers dépasserait la norme minimale de réalisme provisoire de 0,70. Après réflexion, nous faisons l'hypothèse que les élèves de 8<sup>e</sup> année ont surestimé la qualité de

leurs reproductions sur la base qu'ils estimaient que les figures à reproduire étaient banales ou que la tâche de reproduction était facile et élémentaire.

La première question de recherche abordée dans la section portant sur les résultats portait sur l'âge minimal requis pour que le réalisme des élèves puisse être utilisé de façon valide dans un contexte scolaire. Il n'y avait pas d'hypothèse associée à cette question de recherche. Nous cherchions plutôt à déterminer l'âge minimal par observation. Nous avons utilisé le stade de développement cognitif piagétien plutôt que l'âge de l'élève. Cette échelle permet entre autres d'éliminer l'influence des différences individuelles sur le plan du développement cognitif à l'intérieur de chaque groupe d'âge. L'âge ne garantit pas l'atteinte d'un stade de développement cognitif en particulier. En d'autres mots, nous cherchions à déterminer le stade minimal de développement cognitif atteint par les élèves pour que l'utilisation de l'indice de réalisme des élèves (*Rs*) soit pertinente. Nous avons donc compilé l'indice de réalisme (*Rs*) en départageant les garçons des filles et en séparant les deux domaines cognitifs, à savoir le domaine spatio-temporel et le domaine logico-mathématique.

Le premier défi a été d'identifier le seuil correspondant à un point de césure qui discrimine entre le réalisme faible et le réalisme insuffisant. Puisqu'il n'existe pas de normes locales ou de normes pour les enfants, les recommandations provisoires de Leclercq (1993) qui sont validées par la suite par Gilles (1996) ont été retenues. Nous avons donc retenu le point de césure à environ 70 % d'indice de réalisme. Rappelons que les normes de Gilles ont été établies auprès d'étudiants universitaires et que cette population est très différente de notre échantillon. En ce qui concerne l'épreuve des Figures graduées, soit le domaine spatio-temporel, le seuil se situe au niveau de développement OA pour les filles et OB pour les garçons. Nous notons

d'ailleurs un plafonnement de l'indice de réalisme ( $R_s$ ) autour du seuil minimal de réalisme fixé à 0,71. Pour l'épreuve des Concentrations, soit le domaine logico-mathématique, le seuil se situe au niveau IIA pour les filles et quelque part entre IIA et IIB pour les garçons. Nous observons que le stade minimal requis est un peu plus élevé chez les garçons que chez les filles dans les deux domaines. Il semble que le réalisme prenne plus de temps à se développer dans le domaine logico-mathématique pour tous les élèves que dans le domaine spatio-temporel. Il faut dire qu'à moins d'être au début des opérations concrètes, on ne peut pas espérer de réalisme sur les proportions qui est une notion opératoire formelle. Pour l'épreuve des Figures graduées, l'élève nécessite un certain développement au niveau de la fonction symbolique ou de la représentation. Ajoutons également que les notions sur les proportions ne sont pas enseignées ni acquises avant la 6<sup>e</sup> année.

Nous concluons que les élèves en 2<sup>e</sup> année ont presque tous atteint le niveau cognitif OA et OB selon l'épreuve des Figures graduées. L'utilisation des degrés de certitude et de l'indice de réalisme est pertinente pour ces élèves surtout dans des tâches du domaine spatio-temporel. En ce qui concerne le réalisme dans les tâches du domaine spatio-temporel, les recherches ultérieures pourraient cibler des groupes d'élèves plus jeunes de la première année à la quatrième année par exemple. Ceci réduirait le risque de plafonnement et permettrait de déterminer avec plus de précision le stade de développement cognitif minimal pour être considéré réaliste.

Ces résultats soulèvent des questions reliées à la stabilité du trait métacognitif qu'est le réalisme. On constate que le réalisme d'un élève semble varier d'un domaine d'exigence cognitive à l'autre. Nous pouvons certainement faire l'hypothèse que le réalisme d'un élève peut varier d'une matière d'enseignement à l'autre. Nous pouvons également faire l'hypothèse que ce

trait n'est pas stable dans le temps. Si on considère que l'autoévaluation est une compétence qui se développe dans le temps, le réalisme devrait normalement suivre.

La deuxième question de recherche abordée dans la section portant sur les résultats portait sur l'impact potentiel des différences individuelles sur le réalisme des élèves. Les effets principaux sont significatifs pour le sexe et pour le niveau scolaire. Il y a interaction entre le sexe et le niveau scolaire. Nous avons noté antérieurement que les distributions des résultats ne sont pas nécessairement conformes aux postulats de linéarité et d'homoscédasticité. Le test de Box confirme cette observation.

Les analyses subséquentes ont permis de constater une différence significative entre les moyennes de réalisme des filles et des garçons. Le réalisme des élèves est lié au sexe de l'élève dans le domaine spatio-temporel. Les résultats soutiennent également que le réalisme de l'élève varie en fonction de son niveau scolaire. Nous constatons une tendance baissière du réalisme de l'élève en fonction du niveau scolaire. La différence de réalisme entre les garçons et les filles de 2<sup>e</sup> année explique l'interaction observée. Une faible proportion de la variance est expliquée par les différents niveaux de ces conditions. Les différences observées, quoique significatives, sont néanmoins relativement faibles considérant qu'elles se situent autour du seuil minimal de réalisme. Des recherches qui cibleraient des élèves plus jeunes pourraient vérifier à nouveau l'effet du sexe. Le niveau scolaire manque de conformité au postulat d'homoscédasticité de façon significative alors que ce n'est pas le cas pour le sexe des élèves.

Toujours dans la deuxième question de recherche, nous abordons les analyses liées à l'épreuve des Concentrations ou au domaine logico-mathématique. Contrairement à la

distribution des résultats de l'épreuve des Figures graduées, celle de l'épreuve des Concentrations est conforme à tous les postulats et respecte les conditions d'utilisation de l'analyse de la variance. Les résultats soutiennent l'hypothèse que l'âge ou le niveau scolaire peuvent être liés au réalisme des élèves. Les tests de comparaison *a posteriori* montrent des différences significatives entre chaque niveau scolaire. Il y a aussi une différence significative entre les moyennes des filles et des garçons. Les proportions de la variance qui sont expliquées par le sexe et l'interaction sont très faibles. Cette proportion est plus élevée pour le niveau scolaire. À cet égard, Hansson, Rönnlund, Juslin, et Nilsson (2008) avaient relevé une interaction entre l'âge de l'élève, son niveau d'habileté cognitive et son réalisme. Il apparaît donc que le réalisme des élèves est lié directement au sexe de l'élève dans le domaine logico-mathématique. Les résultats soutiennent que le réalisme des élèves varie en fonction de son niveau scolaire, et la tendance est haussière et quasi linéaire avec le niveau scolaire. Chez les plus jeunes, les garçons tendent à être moins réalistes que les filles. Cependant, l'effet expérimental est faible.

Il était attendu que le niveau scolaire, par ricochet le niveau de développement cognitif atteint, ait un impact sur le réalisme. Il est facile de le constater avec les résultats de l'épreuve des Concentrations. Plus l'élève est développé sur le plan cognitif, plus l'élève est réaliste. En ce qui concerne les différences entre garçons et filles, on peut émettre l'hypothèse qu'il semble y avoir un fonctionnement différentiel ou à tout le moins un décalage en fonction du sexe, et ce, surtout chez les plus jeunes. La différence de réalisme entre les garçons et les filles s'observe surtout chez les élèves de deuxième et cinquième année, et ce, dans le cas des deux épreuves. Il apparaît que chez les élèves moins réalistes, les garçons tendent à être moins réalistes que les filles. Un phénomène similaire a été observé antérieurement dans nos recherches sur le réalisme

des adultes (Boulé, 2007). L'interprétation du réalisme des élèves en situation d'évaluation diagnostique pourrait prendre en compte ces dernières observations.

La troisième question de recherche abordée dans la section portant sur les résultats portait sur les domaines cognitifs, à savoir le domaine spatio-temporel et logico-mathématique. Nous avons vérifié si le réalisme des élèves varie en fonction du domaine cognitif. Nous notons une différence significative du réalisme moyen selon le domaine et en fonction du niveau scolaire. Il y a une faible interaction entre le niveau scolaire, le sexe et le domaine cognitif. Il n'y a pas de différence significative entre les filles et les garçons. Une assez forte proportion de la variance est expliquée par le niveau scolaire et le domaine cognitif. Il y a donc une différence significative du réalisme moyen des élèves selon le domaine cognitif de l'épreuve. Les résultats soutiennent que le réalisme de l'élève varie en fonction de son niveau scolaire et non en fonction du sexe en ce qui a trait au domaine cognitif de la tâche.

Le réalisme des élèves est donc dépendant du domaine de la tâche pour la population d'élèves de la 2<sup>e</sup> à la 8<sup>e</sup> année. Notons que cette différence pourrait être attribuable en partie ou en totalité aux modalités de *testing* des deux épreuves. L'épreuve des Figures graduées permet à l'élève de constater visuellement en temps réel la qualité de reproduction de ses dessins, ce qui n'est pas possible dans le cas de l'épreuve des Concentrations portant sur le choix de la proportion ou du ratio le plus élevé. Rappelons également que ces notions de proportion ne sont pas enseignées avant la 6<sup>e</sup> année. L'absence de familiarité avec la tâche peut aussi avoir un effet sur le réalisme. Il y a aussi, tel que décrit antérieurement, le fait que l'épreuve des Concentrations fait appel à des opérations concrètes. L'épreuve des Figures graduées fait appel à la

représentation, à la fonction symbolique, ce que Piaget appelait le stade intuitif d'où un effet de plafonnement pour les élèves de la population étudiée.

La quatrième question de recherche abordée dans la section des résultats portait sur l'écart entre le niveau cognitif atteint par l'élève et l'exigence cognitive de la tâche. Nous avons fait l'hypothèse que l'élève serait plus réaliste lorsque le niveau cognitif atteint par celui-ci est supérieur ou égal à l'exigence cognitive de la tâche. L'élève serait moins réaliste si son niveau de développement est inférieur à l'exigence cognitive de la tâche.

Nous avons comparé le réalisme des élèves dans les tâches qui correspondent à leur niveau opératoire ou qui leur sont inférieures avec le réalisme manifesté dans les tâches qui sont de niveau opératoire supérieur. À titre d'exemple, pour un élève ayant atteint le stade cognitif IIA dans l'épreuve des Concentrations, nous comparons la moyenne du réalisme de cet élève dans les tâches de niveaux IA, IB, IIA avec la moyenne de son réalisme dans les tâches de niveaux IIB, IIIA, IIIA2, IIIB. Les effets expérimentaux sont forts et pertinents pour le sexe et pour l'écart entre le réalisme pour les items d'exigence cognitive au niveau atteint ou inférieur et celles d'exigences supérieures. Il y a interaction entre le sexe et l'écart entre les items d'exigence cognitive au niveau atteint ou inférieur et celles d'exigences supérieures. Puisque les effets expérimentaux sont pertinents, des analyses supplémentaires sont effectuées. Comme noté antérieurement, les distributions des résultats ne sont pas nécessairement conformes aux postulats de linéarité et d'homoscédasticité.

En ce qui concerne l'épreuve des Figures graduées, l'analyse a permis de constater que le réalisme moyen varie en fonction de l'écart entre le degré d'exigence cognitive de la tâche au

niveau cognitif atteint par l'élève et le degré d'exigence cognitive de la tâche qui excède le niveau cognitif atteint par l'élève. Le réalisme de l'élève ne varie pas en fonction du sexe de l'élève. Il n'y a pas d'interaction avec le sexe de l'élève. Les résultats sont plus marquants pour l'épreuve des Concentrations. Le réalisme moyen varie en fonction de l'écart d'exigence cognitive et du sexe de l'élève Il y a interaction avec le sexe de l'élève.

L'analyse a permis de constater que le réalisme moyen varie en fonction de l'écart entre le degré d'exigence cognitive de la tâche au niveau cognitif atteint par l'élève et le degré d'exigence cognitive de la tâche qui excède le niveau cognitif atteint par l'élève. Ce constat corrobore ceux des recherches qui ont porté sur le réalisme et la difficulté des items (Boulé, 2007; Boulé et Laveault, 2011; De Finetti, 1965; Fabre, 1980; Gilles, 1996; Howie et Roebbers, 2006; Jacobs, 1974; Kleitman et Stankov, 2001; Leclercq, 1993; Stankov, 1998). Ces recherches n'ont cependant pas tenu compte de l'exigence cognitive de la tâche ni de la nature curvilinéaire de la relation. Au plan théorique, nos résultats indiquent que le réalisme de l'élève est le plus bas à l'intérieur de la zone proximale de développement, et qu'à l'extérieur extrême de cette zone, l'élève est plus réaliste quant à ce qu'il sait ou ne sait pas vraiment.

La Figure 4 illustre une relation en forme de « U » pour représenter la diminution du réalisme lorsque l'exigence cognitive de la tâche excède le niveau de développement cognitif atteint par l'élève. Nous avons proposé que dans les cas extrêmes d'items exigeants sur le plan cognitif, le réalisme de l'élève augmente puisque l'élève serait certain de ne pas réussir la tâche. On constate que le réalisme tend à remonter lorsque l'exigence cognitive de la tâche est nettement au-dessus du stade cognitif atteint par l'élève. Ce constat soutient que la diminution de réalisme pour les tâches qui excèdent le niveau cognitif atteint par les élèves est curvilinéaire et

probablement en forme de « U ». La recension des écrits n'avait pas relevé d'indications sur l'allure possible de la diminution et de sa courbe. Nous avons anticipé par raisonnement que cette diminution était curvilinéaire. Les données provenant des deux épreuves appuient ce raisonnement théorique. Il s'agit d'une contribution inattendue de cette recherche.

Le cadre conceptuel montrait un algorithme de décision pour les cas où les élèves surestiment la qualité de leurs réponses (voir Figure 3). Les résultats obtenus montrent que l'élève tend à être moins réaliste quand l'exigence cognitive de la tâche excède le niveau développemental atteint par l'élève. Le personnel enseignant peut en tenir compte dans ses choix concernant la régulation de l'enseignement.

## LIMITES ET RECOMMANDATIONS

Cette recherche avance quelques idées originales quant à l'utilisation des degrés de certitude et de l'indice du réalisme dans un contexte diagnostique. Les questions de recherche visaient à examiner certains aspects de la validité de leur utilisation dans un contexte scolaire avec des élèves du cycle primaire et intermédiaire. Pour être utiles et valides, nous devons vérifier si les degrés de certitude et l'indice de réalisme peuvent être employés avec des élèves de tous les âges, de niveaux de développement cognitif différents, et ce, sans égard au sexe des élèves.

Le devis utilisé dans cette recherche présente quelques difficultés qui limitent les généralisations que nous pouvons tirer des résultats. D'une part, nous avons constaté un effet de plafonnement de l'indice de réalisme avec l'épreuve des Figures graduées dans une population d'élèves de la 2<sup>e</sup> à la 8<sup>e</sup> année. Nous avons constaté après coup que les résultats des élèves à cette épreuve nous auraient appris probablement plus si l'échantillon d'élèves avait été constitué d'élèves plus jeunes. Surtout pour le domaine spatio-temporel, des recherches ultérieures pourraient cibler un échantillon d'élèves provenant de la première à la sixième année. Il devrait être assez facile de créer des collaborations de recherche avec des conseils scolaires. Les conseils scolaires cherchent à augmenter leur arsenal d'outils diagnostiques de type non verbaux pour les élèves qui ne maîtrisent pas la langue d'enseignement. De plus, les outils fondés sur les stades développementaux piagétiens intéressent particulièrement les équipes de psychopédagogie des conseils scolaires. Nous notons une augmentation de l'intérêt à mesurer et à obtenir des données probantes sur la réussite des élèves ayant des besoins particuliers.

Les analyses pour les données provenant de l'épreuve des Figures graduées ont été affaiblies par le manque d'homogénéité de la variance. Il en résulte que les coefficients de détermination et la variance expliquée par les variables indépendantes sont faibles. Un échantillon plus ciblé autour de jeunes élèves pourrait probablement régler le problème. Une attention particulière devrait être portée envers l'homogénéité de la variance dans les recherches ultérieures qui utilisent cette épreuve. Notons que nous aurions pu effectuer des transformations sur les distributions afin de créer une plus grande homogénéité de la variance. Ces distributions transformées auraient probablement permis de se conformer davantage aux postulats. Cependant, nous considérons que la pertinence de nos analyses n'aurait pas augmenté en dissimulant l'effet de plafonnement constaté.

Cette recherche met en lumière un effet potentiel des différences individuelles sur le réalisme des élèves. Nous considérons que la différence d'âge et du niveau de développement cognitif atteint ont un impact considérable sur le réalisme des élèves. Des différences entre garçons et filles sont par moment perceptibles surtout lorsqu'on prend le niveau scolaire en considération. De manière à éviter de dissimuler certains effets, le sexe de l'élève devait être pris en considération pour certaines analyses. Nos recherches antérieures avaient montré que cette variable devait être contrôlée statistiquement (Boulé, 2007)

Le réalisme des élèves est différent selon le domaine cognitif de l'épreuve. La nature des tâches des deux épreuves est également très différente. Il n'est pas clair si ces différences sont liées à la nature des tâches et des modalités de réponse ou au domaine cognitif en soi. L'élève qui s'autoévalue avec un dessin concret devant les yeux a un avantage qu'il n'a pas avec l'épreuve des Concentrations. Notons également que l'introduction des concepts de base qui portent sur les

concepts liés aux probabilités et aux pourcentages n'est prescrite par le curriculum de l'Ontario qu'à partir de la 6<sup>e</sup> année. Ceci pourrait avoir un impact sur le degré d'aisance qu'ont les élèves plus jeunes envers les concepts de proportion. À la limite, on peut se demander si les résultats pourraient être différents dans une juridiction où les concepts de proportion sont enseignés à différents moments du parcours scolaire des élèves. Néanmoins, il semble avisé de prendre en considération le domaine cognitif de la tâche quand on compare les recherches qui portent sur le réalisme des élèves.

En ce qui concerne la courbe curvilinéaire de réalisme à mesure que l'exigence cognitive de la tâche s'éloigne du niveau cognitif atteint par l'élève, des recherches ultérieures pourraient tenter de reproduire ces résultats en utilisant des échantillons d'élèves de différents niveaux scolaires. Notons que l'échantillon d'élèves doit être constitué d'élèves assez jeunes. Si on a que des élèves adolescents qui ont déjà atteint le niveau cognitif IIIA, il n'est pas possible d'observer la diminution de réalisme puisque l'épreuve n'a pas de tâches suffisamment éloignées de la zone proximale de développement pour rendre le phénomène observable. De plus, il est possible que la démonstration ne soit pas réalisable considérant que ce n'est qu'un quart à un tiers des élèves selon Noelting (1980a) qui atteignent le niveau formel avec l'épreuve des Concentrations. Les démonstrations présentées utilisent le réalisme moyen des élèves. Des analyses plus pointues peuvent traiter le réalisme sur une base individuelle. De plus, les épreuves pourraient utiliser plus d'items représentatifs de chaque niveau.

Les épreuves de Noelting permettent de déterminer le stade de développement cognitif atteint par l'élève. Nous connaissons également le niveau d'exigence cognitive des tâches. Ces épreuves étaient tout indiquées aux fins de cette recherche. Dans un contexte d'évaluation sur le

terrain dans une matière comme l'histoire, le niveau d'exigence cognitive des tâches n'est pas connu directement. Nous pouvons néanmoins y arriver en développant *a priori* un tableau de spécification qui prend en considération les contenus, mais aussi le niveau d'exigence cognitive des tâches. À titre d'exemple, la taxonomie de Bloom peut être utilisée dans ce contexte.

Les épreuves utilisées ont fait l'objet de plusieurs études de validation dans le passé. Les données recueillies par le conseil scolaire pourraient permettre de reprendre ces efforts de validation. Étant donné la quantité considérable de données recueillies, des approches fondées sur la théorie des réponses aux items (Bond, 2007) pourraient être considérées. Des analyses conformes à la théorie de la généralisabilité pourraient permettre de prendre pour compte l'erreur inhérente liée aux jugements des juges dans les cas de l'épreuve des Figures graduées.

L'échelle de Piaget pose certaines limites par rapport aux analyses statistiques. Il s'agit d'une échelle ordinale qui n'est pas proportionnelle. Nous n'avons pas vraiment d'options. Les recherches ultérieures pourraient en tenir compte.

Cette étude utilise des données secondaires. Elles ont été recueillies par un conseil scolaire dans le cadre d'une étude longitudinale portant sur des outils non verbaux de dépistage. Le conseil a administré jusqu'à maintenant les deux épreuves de Noelting à plus de mille deux cents élèves. Des études supplémentaires seront éventuellement effectuées pour vérifier les liens entre les scores aux épreuves, le réalisme et les résultats à des instruments de mesure. Puisque les données proviennent d'un seul conseil, des recherches ultérieures pourraient tenter de répéter certains aspects de cette recherche en portant une attention particulière à l'échantillonnage.

Les recherches ultérieures pourraient étudier sur le terrain la prise de décision qui découle de l'utilisation de l'algorithme tel que suggéré dans notre cadre conceptuel. La pertinence des interventions au niveau de la régulation de l'enseignement qui découle de la prise de décision par le personnel enseignant devrait aussi être examinée. Cette étude pourrait contribuer à justifier la pertinence du réalisme dans un contexte diagnostique, tant pour l'élève que pour le personnel enseignant. Sur ce dernier point, pour que le réalisme favorise le développement de l'autorégulation chez l'élève, un feedback doit être remis à l'élève. Des recherches ultérieures pourraient examiner le réalisme dans un contexte d'évaluation formative pour les plus jeunes. Leclercq et ses collaborateurs ont examiné le feedback sur le réalisme avec des étudiants adultes et ont proposé plusieurs modèles.

Nos recherches antérieures (Boulé, 2007) faisaient état d'une forte proportion de données manquantes ou de copies incomplètes. De manière à minimiser l'impact de ce phénomène, les élèves ont reçu plusieurs rappels qu'il ne fallait pas oublier d'exprimer leur degré de certitude pour chaque question et pour chaque tâche. Les élèves ont aussi assisté à une démonstration et ont eu une pratique au préalable. Comme Leclercq (1993) le soulignait, nous considérons que les élèves doivent être formés à l'utilisation des degrés de certitude et ils doivent se pratiquer. Une surveillance étroite durant l'administration des épreuves peut réduire considérablement le nombre de copies incomplètes.

Cette recherche visait à permettre de mieux comprendre les facteurs qui affectent le réalisme. Le réalisme de l'élève fait partie du processus d'autoévaluation. Le réalisme de l'élève peut contribuer à une évaluation diagnostique plus valide, fidèle et utile dans certaines conditions. Nous pouvons accroître l'utilité de l'évaluation diagnostique si nous pouvons mieux

comprendre comment les différences individuelles comme le sexe de l'élève, son niveau scolaire et son niveau de développement cognitif atteint peuvent faire varier le réalisme des élèves dans différents domaines.

Les stratégies d'évaluation des ministères de l'Éducation encouragent l'autoévaluation et la mesure de la compétence qu'ont les élèves à réguler leur propre apprentissage. La modalité de réponse qu'est le degré de certitude augmente considérablement la durée de la passation du test. Le concepteur de test doit déterminer la meilleure façon de permettre à l'élève de s'autoévaluer. Il pourrait paraître plus économique de demander à l'élève d'autoévaluer sa performance globale plutôt qu'à chaque question ou chaque tâche. Évidemment, il ne serait pas possible de calculer l'indice de réalisme tel qu'on l'a présenté dans cette recherche si on choisissait cette dernière option. Ajoutons que les degrés de certitude représentent une modalité de réponse supplémentaire qui revêt un certain degré de complexité même si les élèves ontariens sont habitués à s'autoévaluer.

Des tests diagnostiques sont administrés dans le but d'identifier les lacunes individuelles et des carences dans l'enseignement. En plus, nous devons favoriser le développement chez l'enfant de la capacité à s'autoévaluer. Sur le plan pratique, des tests rapides à administrer et efficaces dans ce qu'ils peuvent nous apprendre sont donc nécessaires. L'élaboration de ces instruments doit tenir compte des domaines évalués, des situations de connaissances partielles et des facteurs individuels qui peuvent avoir un impact sur leur fidélité et leur validité.

L'étude du réalisme peut contribuer à mieux comprendre les performances différentielles de différents groupes d'élèves comme les garçons et les filles. Des élèves qui surestiment la

qualité de leurs réponses peuvent négliger de réviser leurs réponses. Aussi, le réalisme pourrait avoir une utilité pour les intervenants qui agissent auprès d'élèves ayant des besoins particuliers qui sont souvent des garçons.

Cette recherche a fait usage d'outils diagnostiques non verbaux, c'est-à-dire que les élèves qui ne maîtrisent pas la langue d'enseignement peuvent néanmoins se prêter à l'exercice comme les autres élèves. Ceci présente un intérêt réel pour les conseils scolaires qui doivent évaluer des élèves nouvellement arrivés au pays ou qui ne maîtrisent pas nécessairement la langue d'enseignement. La capacité à s'autoévaluer des élèves est probablement tributaire jusqu'à un certain point de l'enseignement qu'ont reçu les élèves. Il est fort probable que le développement de la capacité à s'autoévaluer varie considérablement d'un pays et d'une culture à l'autre. Cette variation résulte probablement de la fréquence d'utilisation de l'autoévaluation chez l'élève. Même si les élèves plus jeunes ne semblent pas nécessairement minimalement réalistes, l'autoévaluation s'apprend et on doit commencer plus tôt que plus tard. Il suffit d'en tenir compte lors du développement d'instruments et de l'interprétation des données recueillies.

Finalement, il n'est pas clair que les normes de Gilles (1996) utilisées avec des adultes universitaires s'appliquent à des élèves plus jeunes. Le seuil minimal de réalisme de 0,71 semble élevé. Dans le cas de l'épreuve des Figures graduées, le réalisme moyen des élèves plafonne au niveau du seuil minimal. Pour l'épreuve des Concentrations, la plupart des élèves n'atteignent pas le seuil en question. Ce seuil ne correspond pas au point d'inflexion inférieur de la courbe normale pour la distribution du réalisme des élèves de notre échantillon. Des normes locales pour les élèves plus jeunes devront éventuellement être établies et validées afin de faciliter l'interprétation des résultats des recherches subséquentes. Il faudra déterminer si ce seuil minimal

de réalisme doit tenir compte du niveau de développement cognitif des élèves, l'exigence cognitive des tâches et l'écart entre l'exigence cognitive des tâches et la zone proximale de développement de l'élève. Les normes utilisées dans cette recherche étaient provisoires.

Certaines applications pratiques découlent de cette recherche. Dans un contexte diagnostique, l'information sur le réalisme des élèves peut permettre aux enseignantes et enseignants de réguler leur enseignement. De plus, le personnel enseignant peut prendre en considération l'impact sur le réalisme de l'écart entre le niveau d'exigence cognitive des tâches et la capacité d'assimilation de l'élève. L'âge de l'élève est également un facteur important. Les modalités d'autoévaluation comme les degrés de certitude devraient être introduites tôt dans le parcours scolaire de l'élève. Il faut favoriser l'apprentissage de l'autoévaluation.

Dans un contexte d'évaluation formative, le personnel peut ajouter une forme de *feedback* destiné à l'élève. De cette manière, le réalisme pourrait contribuer au développement de l'autorégulation chez l'élève. Dans la pratique, nous n'avons cependant pas de modèles à proposer et nous n'avons pas d'exemples de pratiques réussies avec les élèves du cycle primaire.

## RÉFÉRENCES

- Allal, L. (1991). Vers une pratique de l'évaluation formative : matériel de formation continue des enseignants. Bruxelles : De Boeck.
- Allal, L. (1999). Impliquer l'apprenant dans le processus d'évaluation : promesses et pièges de l'autoévaluation. Dans C. Depover et B. Noël (Éds.), L'évaluation des compétences et des processus cognitifs (pp. 35-55). Bruxelles : De Boeck.
- Allal, L. (2010). Assessment and the Regulation of Learning. International Encyclopedia of Education, 3, 348-352.
- Allwood, C. M., Innes-Ker, A., Holmgren, J., et Fredin, G. (2008). Children's and adults' realism in their event-recall confidence in response to free recall and focused questions. Psychology, Crime & Law, 14, 529–547.
- Anderson, L.W. (Éd.), Krathwohl, D. R. (Éd.), Airasian, P. W., Cruikshank, K.A., Mayer, R. E., Pintrich, P. R., Raths, J. & Wittrock, M. C. (2001). A taxonomy for learning, teaching, and assessing : A revision of Bloom's Taxonomy of Educational Objectives. New York : Longman.
- Bloom, B.S., Hastings, T., et Madaus, G.F. (1971). Handbook on Formative and Summative Evaluation of Student Learning. New York: McGraw-Hill.
- Bloom, B.S., Madaus, G.F., et Hastings, T. (1981). Evaluation to Improve Learning. New York : McGraw-Hill.
- Bond, T.G. (1995). Piaget and Measurement II: Empirical validation of the Piagetian model. Archives de Psychologie, 63, 155-185.

- Bond, T.G., et Fox, C.M. (2007). *Applying the Rasch Model : Fundamental Measurement in the Human Sciences*. New Jersey : Lawrence Erlbaum Associates.
- Boulé, S. (2007). Utilisation du degré de certitude dans un contexte d'évaluation diagnostique critériée. Thèse de maîtrise non publiée. Canada : Université d'Ottawa.
- Boulé, S., et Laveault, D. (2011). Utilisation du degré de certitude et du degré de réalisme dans un contexte d'évaluation diagnostique. Dans G. Raïche, K. Paquette-Côté et D. Magis (Éds), Des mécanismes pour assurer la validité de l'interprétation de la mesure en éducation : Vol 2-L'évaluation. Montréal : PUQ.
- Burton, R.A. (2007). On being certain. New York : St Martin's Press.
- Butler, A. C., Karpicke, J. D., et Roediger, H. L. (2008). Correcting a Metacognitive Error: Feedback Increases Retention of Low- Confidence Correct Responses. Journal of Experimental Psychology: Learning, Memory, and Cognition, 34(4), 918–928.
- Cardinet, J. (1977). Objectifs éducatifs et évaluation individualisée. Neuchâtel : Institut romand de recherches et de documentation pédagogiques.
- Cardinet, J. (1988). Évaluation scolaire et mesure (2<sup>e</sup> éd.). Bruxelles : De Boeck.
- Cloutier, R., et Girard, M.-A. (1976). Comparaison du développement dans deux épreuves graphiques : Figures graduées et Bender-Gestalt. Enfance, 3, 235-265.
- Crocker, L. et Algina, J. (1986). Introduction to classical and modern test theory. Florida : Harcourt Brace Jonanovich.
- Cronbach, L.J. (1963). Course improvement through evaluation. Teachers College Record, 64(5), 672-683.

- Crooks, T.J. (1988). The Impact of Classroom Evaluation Practices on Students. Review of Educational Research, 58(4), 438-481.
- De Ketele, J.M. (1993). L'évaluation conjugquée en paradigmes. Revue française de pédagogie, 103, 59-80.
- Delcomminette, S., Chapellier, B., et Leclercq, D. (2012). L'évaluation longitudinale par étapes de la compétence (ELEC) à mener une recherche et à en produire le rapport. Actes de colloque : ADMÉÉ-Europe.
- De Landsheere, G. (1979). Dictionnaire de l'évaluation et de la recherche en éducation. Paris : PUF.
- De Finetti, B. (1965). Methods for discriminating levels of partial knowledge concerning a test item. British Journal of Mathematical and Statistical Psychology, 13, 87-123.
- Fabre, J. M. (1980). Jugement et certitude. Berne-Francfort : Lang.
- Flavell, J. H. (1976). Metacognitive aspects of problem solving. Dans L. B. Resnick (Éds.), The nature of intelligence (pp. 231-235). Hillsdale : LEA.
- Gilles, J. L (1996). Utilisation des degrés de certitude et normes de réalisme en situation d'examen et d'auto-estimation à la Faculté de Psychologie et des Sciences de l'Éducation de l'Université de Liège. Actes de colloque : ADMÉÉ-Europe.
- Gilles, J. L. (2002). Qualité spectrale des tests standardisés universitaires. Thèse de doctorat non publiée. Liège : ULG.
- Gilles, J. L. (2005). Le concept de cohérence spectrale appliqué à l'analyse d'items de tests standardisés. Actes de colloque : ADMÉÉ-Europe.

- Grégoire, J. (1999). Que peut apporter la psychologie cognitive à l'évaluation formative et à l'évaluation diagnostique. Dans C. Depover et B. Noël (Éds.), L'évaluation des compétences et des processus cognitifs (pp. 35-55). Bruxelles : De Boeck.
- Guttman, L. (1941). The quantification of a class of attributes: A theory and method of scale construction. Dans P. Horst, P. Wallin et L. Guttman, The prediction of Personal Adjustment. New York: Social Science Research Council.
- Guttman, L. (1950). The basis for scalogram analysis: Measurement and Prediction. New York : Princeton University Press.
- Hadji, C. (1997). L'évaluation démistifiée. Paris : ESF.
- Haladyna, T. M. (2004). Developing and Validationg Multiple-Choice Test Items (3<sup>e</sup> éd.). New Jersey : LEA.
- Hansson, P., Rönnlund, M., Juslin, P., et Nilsson, L.G. (2008). Adult Age Difference in the Realism of Confidence Judgments: Overconfidence, Format Dependence, and Cognitive Predictors. Psychology and Aging, 23(3), 531-544.
- Howie, P., et Roebers, C.M. (2007). Developmental Progression in the Confidence-Accuracy relationship in Event Recall : Insights Provided by a Calibration Perspective. Applied Cognitive Psychology, 21, 871-893.
- Hunt, D. P. (1993). Validity, Reliability and Acuity of Self-Assessment in Educational Testing. Dans D. Leclercq et J. Bruno, J. (Éds.), Item Banking: Interactive Testing and Self-Assessment (pp. 177-189). NATO ASI Series. Heidelberg: Springer Verlag.
- Instituts de recherche en santé du Canada, Conseil de recherches en sciences naturelles et en génie du Canada, Conseil de recherches en sciences humaines du Canada, Énoncé de

- politique des trois Conseils : Éthique de la recherche avec des êtres humains, 1998  
(avec les modifications de 2000, 2002 et 2005).
- Jackson, J. M. (1949). A simple and more rigorous technique for scale analysis. In: A manual of scale analysis, Part II. Montreal: McGill Univ.
- Jacobs, S. S. (1974). Behavior on Objective Tests Under Theoretically Adequate, Inadequate and Unspecified Scoring Rules. ERIC document. Accession Number : ED090312.
- Jans, V., et Leclercq, D. (1997). Metacognitive Realism : a cognitive style or a learning strategy. Educational Psychology, 17(1-2), 101-110.
- Jonnaert, P. J., et Laveault, D. (1994). Évaluation de la familiarité de la tâche : quelle confiance accorder à la perception de l'élève? Revue des sciences de l'éducation, 20(2), 271-291.
- Jonsson, A., & Allwood, C. M. (2003). Stability and variability in the realism of confidence judgments over time, content domain, and gender. Personality and Individual Differences, 34(4), 559-574.
- Keast, A., Brewer, N., et Wells, G.L. (2007). Children's Metacognitive judgments in an eyewitness identification task. Journal of Experimental Child Psychology, 97, 286-314.
- Kelderman, H. (1996). Multi-dimensional Rasch Models for Partial-Credit Scoring. Applied Psychological Measurement, 20, 155-168.
- Kleitman, S., et Stankov, L. (2001). Ecological and Person-Oriented Aspects of Metacognitive Processes in Test-Taking. Applied Cognitive Psychology, 15, 321-341.
- Kleitman, S., et Stankov, L. (2007). Self-confidence and metacognitive processes. Learning and élèveal differences, 17, 161-173.

- Knutsson, J., Allwood, C. M., et Johansson, M. (2011). Child and adult witnesses: the effect of repetition and invitation-probes on free recall and metamemory realism, Metacognition Learning, 6, 213-228.
- Koriat, A. (1993). How do we know that we know? The accessibility account of the feeling of knowing. Psychological Review, 100, 609–639.
- Koriat, A., et Levy-Sadot, R. (2001). The combined contributions of the cue-familiarity and accessibility heuristics to feeling of knowing. Journal of Experimental Psychology: Learning Memory and Cognition, 27, 34–53.
- Kulhavy, R. W., Yekovich, F. R., et Dyer, J. W. (1976). Feedback and Response Confidence. Journal of Educational Psychology, 68(5), 522-528.
- Kulhavy, R. W. (1977). Feedback in written instruction. Review of Educational Research, 47, 211–232.
- Kulhavy, R. W., et Stock, W. A. (1989). Feedback in written instruction: The place of response certitude. Educational Psychology Review, 1, 279–308.
- Kulhavy, R. W., Stock, W. A., Hancock, T. E., Swindell, L. K., et Hammrich, P. L. (1990). Written Feedback : Response Certitude and Durability. Contemporary Educational Psychology, 15, 319-332.
- Latouche-Garneau, M. (1977). Application d'une didactique du dénominateur commun en 6<sup>e</sup> année élémentaire et mesure des effets au moyen d'une épreuve opératoire portant sur la notion de rapport. Thèse non publiée. Québec : Université Laval.
- Laveault, D. (1981). Le passage des opérations concrètes aux opérations formelles: une étude des séquences de développement des notions d'arrangement et de proportion, considérées

dans des univers de tâche et de contenu variés. Thèse de doctorat non publiée,  
Québec : Université Laval.

Laveault, D. (1982). Vérification de la catégorisation effectuée par analyse factorielle. Dans G. Noelting (Éd.), Le développement cognitif et le mécanisme d'équilibration. Chicoutimi : Gaëtan Morin éditeur.

Laveault, D. (1983). La passation des opérations concrètes aux opérations formelles : une étude des séquences de développement des notions de proportion et d'arrangement, considérées dans des univers de tâche et de contenu variés. Revue canadienne de psychologie, 37(4), 521-534.

Laveault, D. (1984). Analyse factorielle. Dans G. Noelting et al. (1984). Figures graduées : Cahier clinique. Centre de psychologie cognitive et développementale. Québec : Université Laval.

Laveault, D. (1998). Autoévaluation et régulation des apprentissages. Dans C. Depover et B. Noël (Éds.), L'évaluation des compétences et des processus cognitifs (pp. 35-55). Bruxelles : De Boeck.

Laveault, D., et Grégoire, J. (2002). Introduction aux théories des tests : en psychologie et en sciences de l'éducation (2<sup>e</sup> éd.). Bruxelles : De Boeck.

Laveault, D. (2003). Modèle de planification et d'amélioration scolaire. Manuscrit inédit.

Laveault, D. (2004a). De la « régulation » au « réglage » : élaboration d'un modèle d'autoévaluation fondé sur l'étude de la régulation des apprentissages. Bruxelles : De Boeck.

- Laveault, D. (2004b). L'évaluation des apprentissages, un jeu de serpents et échelles.  
Conférence de clôture du 24e colloque annuel de l'Association québécoise de pédagogie collégiale. St-Hyacinthe : 10 juin 2004.
- Laveault, D. (2009). L'amélioration de l'efficacité du système éducatif : sur quels indicateurs s'appuyer? Dans X. Dumay et V. Dupriez (Éds.), L'efficacité dans l'enseignement : promesses et zones d'ombre. Bruxelles : De Boeck.
- Laveault, D. (2010). Classroom Assessment in Policy Context : French Sources. International Encyclopedia of Education, 3, 432-437.
- Leclercq, D. (1993). Validity, Reliability and Acuity of Self-Assessment in Educational Testing, Dans D. Leclercq et J. Bruno (Éds.), Item Banking : Interactive Testing and Self-Assessment (113-131). NATO ASI Series. Heidelberg : Springer Verlag.
- Leclercq, D., Boxus, E., Brogniez, P., Wuidar, F., et Lambert, F. (1993). The TASTE Approach: General Implicit Solutions in Multiple Choice Questions, Open Books Exams and Interactive Testing, Dans D. Leclercq et J. Bruno (Éds.), Item Banking : Interactive Testing and Self-Assessment (113-131). NATO ASI Series. Heidelberg : Springer Verlag.
- Leclercq, D. (1998). Pour une pédagogie universitaire de qualité. Sprimont : Mardaga.
- Leclercq, D., Denis, B., Jans, V., Poumay, M., et Gilles, J-L. (1999). L'amphithéâtre électronique. Une application : LQRT-SAFE. Dans D. Leclercq (Éd.), Pour une pédagogie universitaire de qualité. Liège : Mardaga.
- Leclercq, D., et Gilles, J. L. (2001). Techniques de mesure dans l'autoévaluation. Dans G. Figari et M. Achouche (Éds.), L'activité évaluative réinterrogée : regards scolaires et socioprofessionnels. Bruxelles : De Boeck.

- Leclercq, D. (2003). Connaissance partielle, analyse spectrale et métacognition. Dans D. Leclercq (Éd), Diagnostic cognitif et métacognitif au seuil de l'université. Liège : ÉUL.
- Leclercq D., Rinaldi A.M. et Ernould C. (2003). Un questionnaire spectral pour l'évaluation des connaissances chez le patient diabétique, Dans Gagnayre et al. (Éds), L'évaluation de l'Éducation Thérapeutique du Patient, Paris : IPCEM.
- Leclercq, D., et Poumay, M. (2003). Analyses éduométriques et indices métacognitifs appliqués aux questions des 10 check-up MOHICAN : vue d'ensemble. Dans D. Leclercq (Éd), Diagnostic cognitif et métacognitif au seuil de l'université. Liège : ÉUL.
- Leclercq, D., et Poumay, M. (2005). Degrés de certitude : Épistémologie, méthodes et conséquences. Actes de colloque : ADMÉÉ-Europe.
- Leclercq, D. (2006). L'Évolution des QCM. Dans G. Figari et L. Mottier-Lopez (Éds.). Recherche sur l'évaluation en éducation. Paris : L'Harmattan.
- Leclercq, D., et Poumay, M. (2009). La Métacognition. Dans D. Leclercq (Éd.). Psychologie Éducationnelle de l'Adolescent et du jeune Adulte. Liège : ÉUL.
- Leclercq, D. (2009). La connaissance partielle chez le patient : pourquoi et comment la mesurer. Educ Ther Patient, 1(2), 201-212.
- Leclercq, D. (2011). Logiciel Spectral-Certitudes version a. Logiciel. LabSET-IFRES : ULG.
- Legendre, R. (2005). Dictionnaire actuel de l'éducation (3<sup>e</sup> éd.). Montréal : Guérin.
- Legendre-Bergeron, M. F., et Laveault, D. (1980). Lexique de la psychologie du développement de Jean Piaget. Chicoutimi : Gaëtan Morin.
- Lehrer, J. (2009). How We Decide. New York : Harcourt.
- Maril, A., Simons, J. S., Mitchell, J. P., Schwartz, B. L., et Schacter, D. L. (2003). Feeling-of-knowing in episodic memory: an event-related fMRI study. NeuroImage, 18, 827–836.

- Martel, J. (1976). Programme d'entraînement cognitif au raisonnement proportionnel basé sur le développement de structures cognitives. Thèse non publiée. Québec : Université Laval.
- Muraki, E., Carlson, J. E., et Bolt, D. M. (1996). Multidimensional partial credit model: Confirmatory Application. NCME 1996 Conference: San Francisco, CA.
- Noël, B. & et Leclercq, D. (2011). Comment développer des capacités cognitives et métacognitives? Dans Ph. Parmentier (Éd.). Recherches et actions en faveur de la réussite en première année universitaire. Vingt ans de collaboration dans la Commission « Réussite » du Conseil interuniversitaire de la Communauté française de Belgique. Bruxelles : CIUF.
- Noelting, G. (1977). Le constructivisme piagétien et la théorie de l'équilibration illustrés par la construction de la notion de proportion. Philosophiques, 4(2), 145-194.
- Noelting, G. (1980a). The development of proportional reasoning and the ratio concept. Part I: Differentiation of stages. Educational studies in mathematics, 11, 217-153.
- Noelting, G. (1980b). The development of proportional reasoning and the ratio concept. Part II: Problem-structure at successive stages; problem-solving strategies and the mechanism of adaptive restructuring. Educational studies in mathematics, 11, 331-363.
- Noelting, G., et Gagné, R. (1980). The development of proportional reasoning in four different contexts. Proceedings of the Fourth International Conference for the Psychology of Mathematics Education, Berkeley, CA.
- Noelting, G. (1982). Le développement cognitif et le mécanisme de l'équilibration. Québec : Gaëtan Morin.

Noelting, G., Ferland, M., Laveault, D., Noelting, H. E., Punde, C. B., et Tessier, R. (1984).

Figures graduées : Cahier clinique. Centre de psychologie cognitive et développementale. Québec : Université Laval.

Noelting, G., Rousseau, J.-P., Bond, T. J., et Brunel, M.-L. (2000). Can qualitative stage characteristics be revealed quantitatively?, Archives de psychologie, 68, 259-275.

Nunziati, G. (1990). L'enjeu, c'est l'autonomie. Cahiers pédagogiques, 280, 57. OQRE (1999).  
Étude numéro 4, Toronto : Auteur.

Peraya, D., Grenon, V., et Jenni, C. (2011). Questionnaire à choix multiples et travaux de groupe.  
Dans J.-G. Blais et J.-L. Gilles (Éds.). Évaluation des apprentissages et technologies de l'information et de la communication. Le futur à notre porte. Québec : PUL.

Piaget, J. (1930). The Child's conception of physical causality. London : Routledge et Kegan Paul.

Piaget, J. (1972). Problèmes de psychologie génétique. Paris : Denoël.

Piaget, J. (1974). La Prise de Conscience. Paris: PUF.

Piaget, J. (1977). Recherches sur l'abstraction réfléchissante, tome 1 : L'abstraction des relations logico-arithmétiques. Paris: PUF.

Piaget, J., et Inhelder, B. (1948). La représentation de l'espace chez l'élève. Paris: PUF.

Pintrich, P.R. (2002). The Role of Metacognitive Knowledge in Learning, Teaching, and Assessing. Theory into Practice, 41(4), 219-225.

Scallon, G. (2000). L'évaluation formative. Québec : ÉRPI.

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. Contemporary Educational Psychology, 19, 460-475.

- Scriven, M. (1967). The methodology of evaluation. Dans: Perspectives of Curriculum Evaluation, 39-83. Chicago : Rand McNally.
- Spearman, C. (1907). Demonstration of formulae for true measurement of correlation. American Journal of Psychology, 18, 161-169.
- Stankov, L. (1998). Calibration curves, scatterplots and the distinction between general knowledge and perceptual tasks. Learning and Individual Differences, 10(1), 29-50.
- Stankov, L., et Lee, J. (2008). Confidence and cognitive test performance. Journal of Educational Psychology, 100(4), 961-976.
- Stankov, L., Lee, J., et Paek, I. (2009). Realism and confidence judgments. European Journal of Psychological Assessment, 25(2), 123-130.
- Stiggins, R. (2004). New Assessment Beliefs for a New School Mission. Phi Delta Kappan, 22-27.
- Vygotsky, L. S. (1986). Thought and language. MA: MIT Press.
- Wegmuller, E. (2002, sept.). L'autoévaluation pour développer l'autonomie. Résonances, 1, 10-12.
- Yao, L., et Schwarz, R.D. (2005). A Multidimensional Partial Credit Model With Associated Item and Test Statistics. AERA 2005 Conference: Montréal.

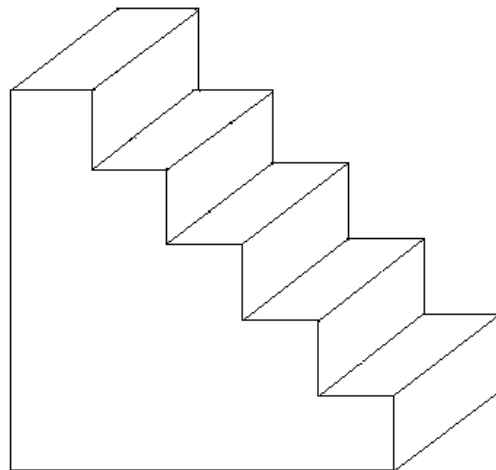
## APPENDICE A

EXEMPLE DE FIGURE GRADUÉE ET DE DIRECTIVES DESTINÉES À L'ÉLÈVE

Voici un exemple de figure graduée qui doit être reproduite. Le cahier clinique de Noelting et al. (1984) présente des exemples types de reproductions réussies et de cas rejetés. Chaque figure est accompagnée de critères descriptifs d'appréciation qui sont utilisés par les juges lors de la session de notation.

Directives verbales et écrites destinées à l'élève lors de la passation du test :

*Reproduis cette figure dans ton cahier sur une nouvelle page. Tu peux effacer ou recommencer sur une autre page au besoin.*



APPENDICE B



EXEMPLE PROVENANT DE L'ÉPREUVE DES CONCENTRATIONS

Voici un exemple d'item de l'épreuve des Concentrations. Il s'agit d'une tâche de type (2,3) vs (1,2). La demande cognitive associée à cette tâche en particulier correspond au stade IIIA<sub>1</sub>, c'est-à-dire, au stade opératoire formel inférieur. Lors de la passation de l'épreuve, une démonstration avec du vrai jus aide à mettre les élèves en contexte authentique. La bonne réponse dans ce cas-ci est A.

Directives verbales et écrites destinées à l'élève lors de la passation du test :

Voici des verres de jus  et des verres d'eau .

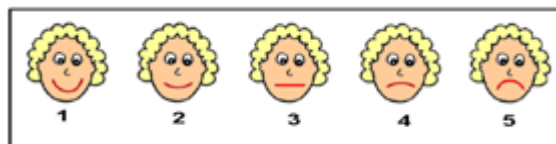
*Quel mélange te donnera un jus plus concentré?*

<b>A</b>		<b>B</b>
		
+	=	+
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

## APPENDICE C

### MODALITÉ PERMETTANT L'EXPRESSION DU DEGRÉ DE CERTITUDE

(À des fins de calcul, les pictogrammes de 1 à 5 correspondent aux valeurs de certitude suivantes : 0,98, 0,74, 0,50, 0,26, 0,02 respectivement)



(À des fins de calcul, les pictogrammes de 1 à 5 correspondent aux valeurs de certitude suivantes : 0,98, 0,74, 0,50, 0,26, 0,02 respectivement)

APPENDICE D

LE CALCUL DU RÉALISME

Le calcul du réalisme de l'élève ( $R_s$ )

$$\bar{e} = \frac{\sum_{i=1}^n (|C_i - T_i| \cdot U_i)}{J_i}$$

$\bar{e}$  = erreur moyenne absolue de certitude

$i$  = indice des degrés de certitude

$n$  = nombre des degrés de certitude

$C_i$  = valeur de la certitude  $i$  (l'échelle doit permettre de l'exprimer en pourcentages)

$R_i$  = nombre de réponses correctes pour la certitude  $i$

$U_i$  = nombre d'utilisation de la certitude  $i$

$T_i$  = taux d'exactitude de la certitude  $i$  (en pourcentages) =  $(R_i / U_i) \times 100$

$J$  = nombre total de réponses =  $\sum_i U_i$

Et l'indice de réalisme ( $R_s$ ) sur une échelle de 0 % à 100 % correspond à :

$$RÉALISME(R_s) = 100 - (\bar{e})$$

### Le calcul de l'indice de réalisation des prédictions par question ( $Rq$ )

Cette équation tirée de Gilles (2002) est similaire à celle utilisée pour calculer le réalisme des élèves ( $Rs$ ). Cependant, nous utilisons les certitudes associées aux questions plutôt qu'aux élèves.

$$\bar{e}_q = \frac{\sum_{i=1}^n (|C_i - T_i| \cdot U_i)}{J_q}$$

$\bar{e}$  = erreur moyenne absolue de certitude

$i$  = indice des degrés de certitude

$n$  = nombre des degrés de certitude

$C_i$  = valeur de la certitude  $i$  (l'échelle doit permettre de l'exprimer en pourcentages)

$R_i$  = nombre de réponses correctes pour la certitude  $i$

$U_i$  = nombre d'utilisation de la certitude  $i$

$T_i$  = taux d'exactitude de la certitude  $i$  (en pourcentages) =  $(R_i / U_i) \times 100$

$J_q$  = nombre total de réponses à la question  $q$  =  $\sum_i U_i$

Et l'indice de réalisme ( $Rq$ ) sur une échelle de 0 % à 100 % correspond à :

$$INDICE\_DE\_RÉALISATION\_DES\_PRÉDICTIONS(Rq) = 100 - (\bar{e}_q)$$